APPLICATIONS OF LOW-COST VOICE RECORD/PLAYBACK TECHNOLOGY IN THE DEVELOPMENT OF SOPHISTICATED AAC AIDS: A CASE STUDY

M. Gradim Gericota[1], M. Lourdes Lourenço[2], J. M. Martins Ferreira[3]
[1]INESC / ISEP [2]CRPCP [3]INESC / UP

ABSTRACT

Voice-output augmentative and alternative communication (AAC) aids play a leading role in the integration of people with speech impairments, but their cost is still a major obstacle preventing many potential users to benefit from this technology.

This paper presents a prototype voice-output AAC aid based on general-purpose voice record/playback technology, where one of the major objectives to achieve consisted of retaining the performance of common commercial systems, while minimising the cost. Three main user interface access methods were implemented and are also described.

INTRODUCTION

The provision of voice-output in AAC aids significantly contributes to an improved interaction capability, but the cost of these devices is presently a barrier for many potential users. It is however interesting to refer that the cost of general purpose voice record/playback technology has steadily dropped over the last years, while at the same time the performance of these modules has steadily increased.

Three are three main reasons that led to the availability of sophisticated low-cost voice record/playback modules:

1. Advances in speech coding technology, important because of the huge cellular communications market, led to complex algorithms enabling low-rate transmission (and therefore simplified storage) of speech signals.
2. Advances in microelectronics technology enabled these complex algorithms to be implemented in increasingly smaller integrated circuits, instead of requiring expensive computing platforms.
3. Finally, the two previous factors combined led to a market opportunity for voice-output devices in many areas, which fostered the appearance of general purpose, very low-cost voice-output modules [1].

These same reasons open excellent opportunities for the development of sophisticated voice-output AAC aids using technology developed for the much wider consumer electronics market. Following a brief description of the main user requirements which were identified by a multi-disciplinary team including therapists and electronics engineers, this paper presents a prototype device based on a very low-cost module readily available, as well as the three main interface methods which were implemented.

USER REQUIREMENTS

Based on the experience of therapists using voice-output AAC systems at the Cerebral Palsy Rehabilitation Centre of Porto, a basic set of user requirements was compiled to drive the technical development phase. The main issues related to the user requirements which were identified can be summarised as follows:

1. The performance of the systems currently in use should be retained. These systems provide in the range of tens of messages, built-in recording capability and direct or scanning access methods.
2. Portability was a desirable issue, since some of the systems in use were essentially meant for desktop usage. additional features related to portability, such as the use of batteries and their autonomy, were therefore important as well.
3. The use of voice-output AAC aids involves a training period where the user regularly improves his skills. The devices should therefore support this evolutionary

acquaintance process, namely through a range of user interface alternatives.

4. Finally, minimum cost is obviously a major requirement.

A technical specification was built upon these requirements, leading to the prototype solution which will now be described in the two following sections.

A "SPEECH STIK" -BASED AAC DEVICE

A survey of voice record/playback systems available on the market led to the choice of one module manufactured by Dallas Semiconductor, the DS2271 "Speech Stik" [2]. This is a highly compact module which includes all the required circuitry in a double-side 4"x1" surface-mount board, from the analog input and output sections to the dedicated blocks for processing the digitised speech signal. Up to 254 messages can be recorded and played back.

The extra circuitry required was therefore restricted to the addition of extra non-volatile memory (to maximise the recording time) and of a dedicated microcontroller to support the required user interface alternatives. An early version of the laboratory prototype is shown in figure 1.

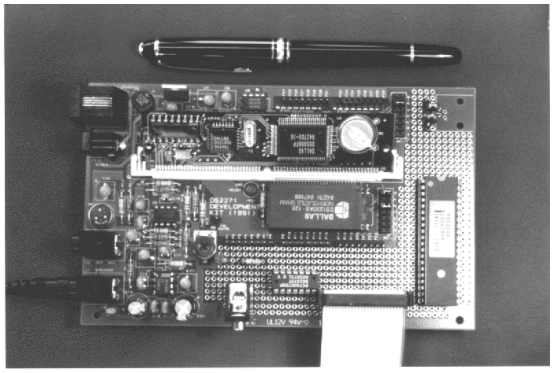

Figure 1: An early laboratory prototype.

USER INTERFACE METHODS

According to the requirements identified at the beginning of the project, three user interface methods were implemented:

1. Direct access through a matrix keyboard.
2. Scanning access through an LED matrix.
3. Scanning access through numerical selection.

It is important to refer that the three interface methods were programmed into the built-in memory (EPROM) of the interface microcontroller, which means that changing from one method to another can be done simply by plugging the required interface into the proper connector of the AAC device.

KEYBOARD ACCESS

A matrix keyboard provides direct access to message selection, with 30 keys allocated to specific messages and 2 keys allocated to control functions.

Since up to 8 pages (8x30=240 messages) are available, one of the control keys implements a "next-page" function. The other control key is used to signal the system that a number of messages/words is going to be selected, and that only after this selection the complete sequence should be output.

In those cases where less experienced users require a smaller number of messages, neighbouring keys can be grouped together under the same message. Four operating modes were implemented and are illustrated in figure 2. Whenever the user selects any key of one particular group, the corresponding message to that group is output.

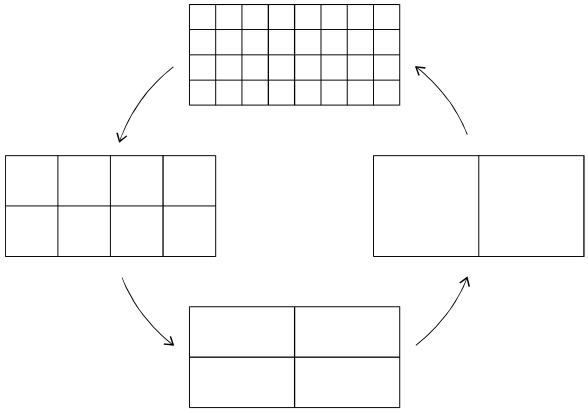

Figure 2: Grouping neighbouring keys.

SCANNING ACCESS THROUGH AN LED MATRIX

A 32-cell with a similar layout to the matrix keyboard can also be used, the main difference

here being that one LED replaces each key. One of several scanning methods can then be selected (linear scanning, row/column scanning, etc.) and the user has only to signal the desired cell by pressing a switch when the respective LED is on. The page control "key" (LED) is again made available to allow the user to select up to 8 pages of 30 messages and the "long phrase" option is retained as well.

SCANNING WITH NUMERICAL SELECTION

In order to meet the portability requirements, a two-digit scanning selection procedure was also made available. This same access method had already been implemented in a previous project [3] and proved to be a useful and simple interface to select up to 100 messages.

Each two-digit code corresponds to one message and the user has one switch to indicate the desired code. The selection process starts with the left digit cycling from 0 to 9. Pressing the switch will stop the left digit in its present value and start the cycling procedure in the right digit. When the user presses the switch again a unique two-digit code has been selected and the corresponding message is then output.

EVALUATION

An evaluation phase is under way at the time of writing at the Cerebral Palsy Rehabilitation Centre of Porto. The main evaluation parameters which are being assessed are as follows:

1. Improvement in interaction capability (conversation has become easier? the number of user-initiated conversation cycles has increased?)
2. Improvement in linguistic capability (syntactic structure has improved? is there now a better use of language?)
3. Communication between peers (do users now talk to one another? has turn taking and conversation time improved?)

The results so far indicate that all the user requirements described were successfully met, but more time is required to gather numerical evidence concerning the evaluation parameters referred above.

CONCLUSION

The work done so far has proved that it is possible to use very low-cost, widely available voice record / play back modules, developed for multi-purpose consumer electronics applications, as a basis for the development of sophisticated voice-output AAC aids. Since the current technological trend is towards lower cost and improved performance, it seems possible to develop voice-output AAC systems where cost will no longer be a barrier.

Further research should still be carried out to identify what implications a universal access to voice-output devices could have in the general field of AAC, namely for user requirements other than those addressed in this work.

REFERENCES

[1] J. Gallant, "Speech synthesis and recognition chips personalise consumer products," *EDN Europe*, pp. 51-57, October 20, 1994.

[2] DS2271 Speech Stik Data Sheet, *Dallas Semiconductor Teleservicing Design Handbook*, 1991, pp. 212-232.

[3] J. M. Ferreira, A. C. Teixeira, J. L. Ramalho, M. L. Lourenço, "A Portable Device to Improve the Communication Ability of People with Cerebral Palsy," *Proceedings of the 6th ISAAC Conference*, Maastricht, NL, pp. 446--448, October 1994.

Manuel Gradim Gericota (mgg@inescn.pt)

INESC - Pr. da República, 93 - 4050 Porto

Tel. 02-2094243 / Fax: 2084172