PROCEEDINGS

RESNA '95 RECREAbility

THE COST X PERFORMANCE TRADEOFF FOR VOICE OUTPUT AAC AIDS: IS IT REAL?

J. M. Martins Ferreira, M. Gradim Gericota, M. Lourdes Lourenço
INESC / University of Porto and Cerebral Palsy Rehabilitation Center
Pr. da República, 93 - 4000 Porto - PORTUGAL

ABSTRACT

The cost of voice output AAC aids is still too high, which prevents many persons with speech impairments to benefit from this technology. However, advances in speech coding techniques and in microelectronics (integrated circuit complexity, and packaging and mounting technology), have enabled new solutions which will in the short future provide a generalized access to voice output AAC aids. This paper presents the basic issues dealing with the technology of voice-output communication aids, describes two new types of solutions based on very low-cost technology and highlights the implications which can be brought to the AAC field itself if a truly generalized access to high-complexity and low-cost voice output devices effectively becomes real.

BACKGROUND

Augmentative and Alternative Communication (AAC) may be defined as "an area of clinical practice that attempts to compensate (temporarily or permanently) for the impairment and disability patterns of individuals with severe expressive communication disorders (i.e., the severely speech-language and writing impaired)" [1].

Both graphical and gestural mode communication systems are extensively used in AAC. Graphical mode systems include such alternatives as PCS (Picture Communication Symbols), PIC (Pictogram Ideogram Communication), Blissymbols, etc., and can be used either with direct selection or scanning techniques. While non-electronic methods for displaying symbols (e.g. communication boards) are basically inexpensive, electronic methods present several advantages. Beside making it much easier for the listener to follow a conversation, electronic communication aids with voice output allow the user to express him/herself even if the listeners are not momentarily paying attention.

Voice output AAC aids can be used with any the previously referred communication systems, although other alternatives are also available. The Minspeak™ system [2] is a good example of a powerful language representation system which would be much less appealing if it didn't have voice output. Both synthesized (where each word is generated by concatenating basic phonemes) and digitized voice output are possible. While the former is usually employed in more powerful systems (where the number of sentences which can be reproduced is not limited), the latter usually produces better quality speech (words or sentences are digitized and stored in digitized form).

Voice output AAC aids make extensive use of speech coding technology, which deals with the compact representation of voice signals for the purpose of efficient transmission or storage [3]. Very substantial progress was made in this area during the past decade, the main reason being its importance for civilian and military communications and for computer related voice applications. Research in this area became very active during World War II due to its strategic importance for encrypted speech communication. Many advanced algorithms for efficient speech coding are now practical due to developments in microelectronics technology, which presently enables the production of ICs (integrated circuits) with more than 4 million transistors.

The importance of speech coding for many areas of the consumer electronics market is responsible for the availability of very sophisticated multi-purpose components and sub-systems, which can also be used to the advantage of people with speech impairments. Among other possibilities, ADPCM (Adaptive Differential Pulse Code Modulation) is one technology for which many dedicated ICs and even pre-assembled modules are now inexpensive and available on the market. The basic idea behind ADPCM consists of sampling the analog speech signal and converting it to digital form, but storing only the difference between the amplitudes of the present and the previous samples. The step size and the correlation factor with previous samples are allowed to adapt to the time-varying statistic properties of the speech signal.

When each analog speech sample is taken with a 12-bit resolution (differentiating among 2^{12}=4096 discrete values), the difference between adjacent samples can be stored using only 4 bits per sample (reducing the required memory to one third, when compared to the full binary representation).

Although digitally coded speech allows transmission rates well within the available bandwidths and provides better noise immunity for communication over long distances, the

compression ratios achievable are not yet entirely satisfactory for storage purposes. ADPCM provides good quality speech transmission at 32 kbps (kilobits per second), but if its digital representation has to be stored, approximately 1 Mbyte of memory is required for 30 seconds of speech.

An alternative technique, which emerged in the last 3 to 4 years, uses standard digital memory cells to directly store and read information in analog form [4]. DAST™ (Direct Analog Storage Technology) significantly improves the storage efficiency when compared to digital coding methods (achieving a 4:1 storage advantage over ADPCM) and is therefore a serious low-cost competitor to ADPCM for many voice record and playback systems.

STATEMENT OF THE PROBLEM

Since speech coding is now a mature technology, it is somehow surprising why a generalized access to voice output AAC aids is not yet possible, specially because one major reason for this fact is that cost is still too high for many potential users. This is indeed the main problem presently restricting access to voice output AAC aids at the Cerebral Palsy Rehabilitation Center of Porto, where a much larger number of persons with speech impairments would like to use this technology.

The aim of the work described in this paper was therefore to use very low-cost technology to develop good quality voice output AAC aids, adaptable to the dynamic requirements of AAC interventions in children and young adults with cerebral palsy.

APPROACH

Following an user requirements identification phase where the main functional and interface issues were agreed, several alternative technologies were considered as possible candidates. It must be said that most of the technologies which were considered were generally able to meet the cost and performance requirements. A first prototype was developed which used a commercially available dedicated ADPCM chip to provide a portable communication device able to store up to 100 messages [5]. The feedback obtained from the users testing this prototype was very positive and it is now in the process of being industrialized. Two additional types of solutions were meanwhile developed, which are described in the following paragraphs.

One of these solutions consists of small and very low-cost prototypes meant to stimulate early communication and language capability acquisition in children with cerebral palsy. This prototype was based on DAST™ technology and is illustrated in figure 1.

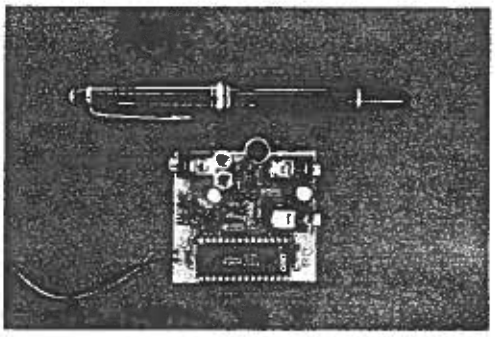

Fig. 1: A voice output device for children with CP

Up to 16 seconds of speech can be stored in this device (speech recording is done by speaking to the microphone shown in figure 1, while the record button is pressed) and the quality of the voice output is perfectly acceptable[1].

The second type of solution is illustrated in figure 2 and consists of more sophisticated devices, where up to 254 messages can be made available and several operating modes selected. Again, recording can be done directly with a built-in microphone, which is common to many of the communication systems already commercially available.

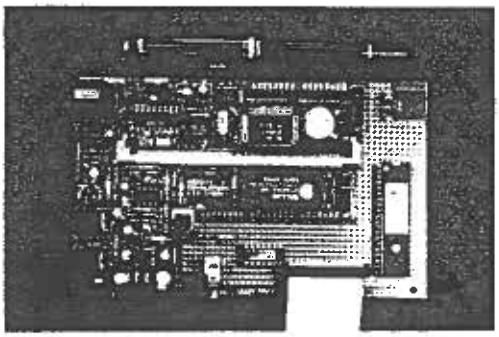

Fig. 2: A DS2271[2]-based voice output device.

[1] In technical terms, the device specifications guarantee a signal-to-noise ratio of 40 dB, a 3-dB bandwidth of 3,4 Khz and a total harmonic distortion of 2% at 1 KHz.

[2] The DS2271 is a multi-purpose speech stick marketed by Dallas Semiconductor.

The main innovation in this prototype is in fact its much lower cost, when compared to the products available on the market, which was made possible by the use of a multi-purpose solution integrating all the required electronics in a very compact module [6]. Several user interface alternatives and operating modes have been implemented. Voice output can follow each selection, but a complete sentence can also be defined before the corresponding phrase is output. Dynamic memory management ensures that all the existing memory can be used even when the words or sentences have different lengths (when a message is to be modified, it is first deleted, then all the following messages are shifted back to fill the hole left by the message removed and only then the new message is stored).

IMPLICATIONS

Within the limited scope of the Cerebral Palsy Rehabilitation Center of Porto, the very low-cost of the range of solutions which were developed in the course of this project met the objective of enabling a generalized access to voice output AAC aids. An immediate implication of this result is that the speech therapists working at this institution are now introducing more sophisticated AAC aids at a much earlier stage, both for children and for young adults. At the same time, a specific program has started to assess the improvement in communication and language acquisition capability, when compared to the use of non-electronic AAC aids.

The identification of additional applications where voice record and playback devices might be of interest to people with speech impairments, specially if very low-cost allows a non-restricted number of devices to be used anywhere, has also started to be considered. Although the evaluation work has so far been related to the specific requirements of people with cerebral palsy, aphasia is an area which will soon be considered as well.

DISCUSSION

The results obtained clearly indicate that speech coding and microelectronics technologies have matured enough so that very sophisticated building blocks are now widely available at inexpensive prices. One major consequence of this fact is that there is really no reason to compromise quality to cost, since the latter is basically dependent on other factors than the cost of the components or sub-systems used in many voice output AAC aids.

On the other hand, it is highly probable that voice output will start to be combined with voice recognition in a larger number of application areas, such as hand-held personal digital assistants, digital answering machines and a variety of voice organizers, such as memos, calendars and to-do lists [7]. The general trend in this field is therefore that cost will continue to decrease while product performance will continue to increase, mainly because additional application areas in the consumer electronics mass market will start to become practical as technology evolves.

REFERENCES

[1] D. Beukelman, P. Mirenda, *Augmentative and Alternative Communication*, Paul H. Brookes Publishing Co. Inc., 1992 (AAC definition according to the American Speech-Language-Hearing Association — ASHA).

[2] G. Van Tatenhove, *What is Minspeak™?*, Prentke Romich Company, 1993.

[3] A. Spanias, "Speech Coding: A Tutorial Review," in *Proc. of the IEEE*, Vol. 82, n° 10, pp. 1541-1582, October 1994.

[4] F. Goodenough, "IC Holds 16 Seconds of Audio Without Power," in *Electronic Design*, January 31, 1991.

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

[6] DS2271 Speech Stick Data Sheet, Dallas Semiconductor Teleservicing Design Handbook, pp. 212-232, July 1991.

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

ACKNOWLEDGEMENTS

This work was partially supported by the Secretariado Nacional de Reabilitação through an HORIZON project entitled "Implementação de um Laboratório de Informática e de Electrónica".

José Manuel Martins Ferreira
INESC / University of Porto
Pr. República, 93 - 4000 Porto - PORTUGAL
Tel. 351-2-2094242 - Fax: 351-2-2084172
E-mail: jmmf@inescn.pt