AIDS FOR THE HANDICAPPED BASED ON "SYNTE 2" SPEECH SYNTHESIZER

M.A. Karjalainen, U.K. Laine, and R.O. Toivonen

Tampere University of Technology
Tampere, Finland

ABSTRACT

SYNTE 2 is a low-cost, high-quality, text-to-speech synthesizer designed for Finnish but applicable also to other languages if "phoneme writing" is used. After its first presentation in 1977 it has been adapted to many communication aids for the handicapped. The first application was a portable speaking machine with unlimited vocabulary for the speech impared. This paper describes the present applications of SYNTE 2, including the speaking machine, a talking data terminal for blind computer programmers, a system for automatic production of spoken information for the blind, etc.

INTRODUCTION

SYNTE 2 is a real text-to-speech synthesizer for Finnish designed to meet a wide variety of applications (1). It consists of a microprocessor (Motorola 6800 + 2 kbyte ROM + 256 byte RAM), a group of special D/A-converters to generate analog control signals, and an analog signal processing part for sound generation (Fig. 1). The latter one is built around a vocal tract model, which is a combination of cascaded and serial types. The synthesizer contains four circuit boards each of size 100 x 160 mm.

SYNTE 2 can be controlled by ASCII-coded text. A sentence is written in, letter by letter, until the end character (.) is given. This starts the real-time synthesis process. If desired, a number of special characters can be used to control speech rate, intonation, phoneme variations, etc.

SYNTE 2 is in serial production. The main idea in its development was to realize aids for communication handicapped (2). It has also found a number of other applications e.g. in audiometry (3) and man-machine communications such as speech response systems via telephone.

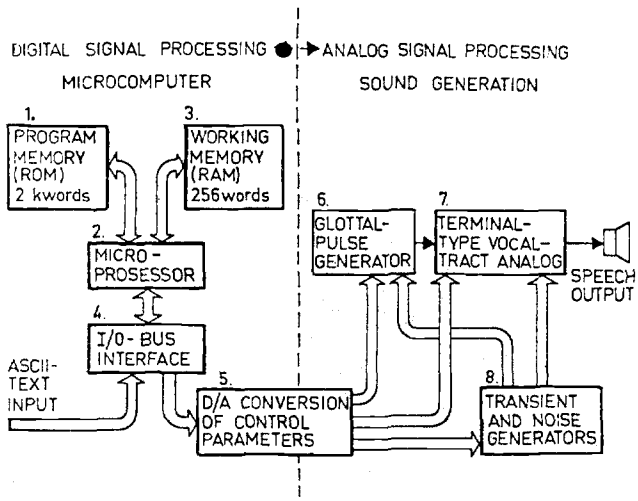

Fig. 1. Block diagram of SYNTE 2 speech synthesizer

SPEAKING MACHINE FOR THE SPEECH IMPARED

Because of the small size and low weight, SYNTE 2 is suitable for a portable and battery-operated speaking machine for the speech impared. The prototype of the device, first time presented in 1977 (4), consists of SYNTE 2, a small keyboard (32 characters), rechargeable batteries, and a loudspeaker in a 75 x 170 x 230 mm enclosure with total weight of about 2 kg (Fig. 2). The effective operation time with full-charged batteries is about 2 hours and the device is equipped with an automatic power on-off switching for easy use. This device can be used as a speaking machine for the speech impared, as a rehabilitation aid for the motory handicapped or as an aid in learning correct writing.

The preliminary tests on SYNTE 2 as used by speech disabled persons have been encouraging. The device has proved to be an useful aid in everyday life for those who have the ability to write with a keyboard. The understandability of single words without context is found to be about 90 %, which corresponds to about 100 % in

CH1559-4/80/0000-0851$00.75 © 1980 IEEE

meaningful sentence context. The voice is natural enough and not tiresome to be listened to.

The most severe disadvantage of the device is the slow speed of keyboard input. The speed after a short learning period (from 3 days to 1 month) is about 1...3 characters per second. This makes it tedious to wait for sentences of length more than four to five words. The users have also found the weight too high and the size little impractical if it is to be carried continuously. Further improvements of the speaking machine are to be realized.

Fig. 2. Prototype of the speaking machine for the speech impared

TALKING DATA TERMINAL

The talking data terminal for blind computer programmers (Fig. 3) is a microprocessor-based (Motorola M6800) intelligent terminal with an ASCII-keyboard, SYNTE 2 speech synthesizer, loudspeaker, headphones, and interfaces to computer or modem, TV-set and line printer. An audio cassette memory and video display with large characters will be added later.

The software to control the functions of the terminal contains many intelligent features. The user can choose different speech output modes by keyboard commands. There are independent definitions to control the modes for the text from computer and from keyboard. The basic modes are: speech output on or off, words to be spoken by spelling or reading, numbers as strings of digits or as whole numbers, and special characters always character by character. Redundant strings of special characters from computer can be defined to be spoken in compact form, e.g. ++++ as "four plus signs".

To make it easier to listen to the speech output, there is a possibility for the user to define a table-coding dictionary, which can be utilized to substitute difficult input strings by new and more plain ones; e.g., LDA → "Load to accumulator" or LF CR → SP. This dictionary is a list of character string pairs in the RAM memory of the terminal and it can be written in, completed, changed or deleted by keyboard commands. The terminal checks, disregards and reports by speech all ill-formed or illegal definitions.

The terminal can further be commanded to use half- or full-duplex in data communication. The speech rate can also be increased or decreased in 17 % steps. In the future the user-selectable modes and the table-coding dictionaries can be stored in a tape cassette and then easily recalled.

Besides, the present simple "speech listing mode" of files, a searching mode will be added, where additional control buttons can be used to move a "speech cursor" forwards and backwards in the file space. A program called "talking text editor" is also planned.

Field tests of the talking data terminal during 9 months have proved the device very useful for blind programmers. The working speed has doubled or increased even more; still no tiring effects or extra stress are found. Besides the so-called Braille terminal the new talking terminal means real progress for blind workers in many "information occupations".

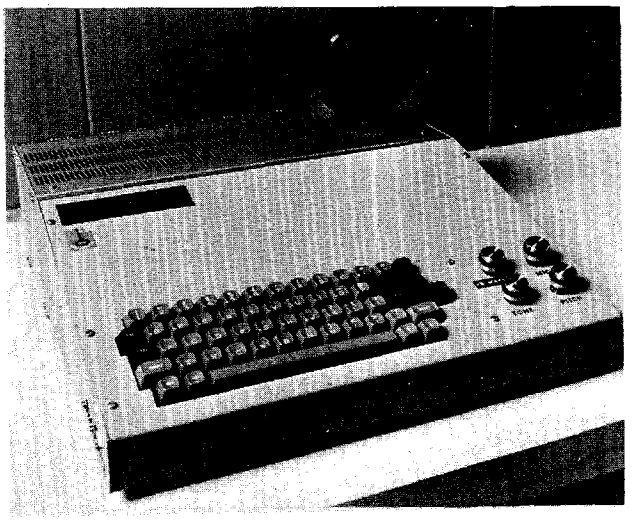

Fig. 3. The talking data terminal

AUTOMATIC PRODUCTION OF SPOKEN INFORMATION MATERIAL FOR THE BLIND

The development of a system for automatic production of spoken information material for the blind by synthetic speech started in 1977. The main aim is to produce "voice books and newsletters" for the blind as tape cassettes from computer-controlled typesetting material.

The prototype system (Fig. 4) consists of a microcomputer (EXORterm 220), dual-drive floppy disk memory (EXORdisk II), talking data terminal described earlier, audio cassette recorder, video display, paper tape reader, and printer. This complex will later be replaced by a more compact and less expensive processor unit (Motorola EXORset 30), a SYNTE 2 synthesizer, and some peripheral devices.

The main part of the work has been the development of programs for text processing. This contains about 10 000 assembly-language instructions to convert arbitrary Finnish text of typesetting material into strings of letters suitable for feeding SYNTE 2 synthesizer. The problem has been to expand abbreviations, numbers, and strings of special characters, to find the boundaries of sentences, to remove control characters for typesetting or to utilize the information carried by them, and to interpret the context-dependent meanings of some special characters.

An interpreting high-level command language is created to make it easy to define different forms of text processing. The basic commands control data transfers and text transformations. Command programs can be written in from keyboard or loaded from floppy disk.

The text is processed in ring-like data structures, which are well suited for continuous input and output flow of characters. There is one ring for command programs to be interpreted and 15 rings for texts to be processed.

The production of spoken information material begins with the reading of typesetting tapes into disk files. Each file will then be transferred in smaller segments into the ring structures, where the text is processed and finally sent to the talking data terminal for video display and voice output. The whole process works automatically when the command program is started. A cassette recorder can be used to store the speech output. Also a blind person is able to use the system because the talking data terminal reports every act of the user and the responses of the system.

The prototype system works well; evaluations of the performence for different kinds of text material are to be done and improvements of details continue. The system is also a good starting point when designing other intelligent speech response systems such as talking data bases and information services, automatic announcement systems, etc.

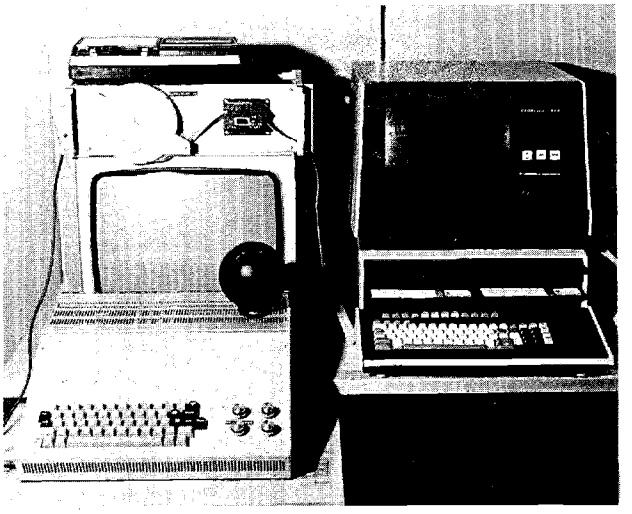

Fig. 4. Prototype system for automatic production of spoken information material for the blind

OTHER APPLICATIONS

The versatility of a text-to-speech synthesizer makes it easy to find a large number of potential applications. One idea under realization is a "discussion machine". It is to be used as an aid for teaching language to deaf children. It consists of two communication channels by radio; one from a small keyboard of the teacher to an alphanumeric text panel of the deaf child, the other from the keyboard of the child to a speech synthesizer of the teacher. The aim is to make the system portable and to study the increase in conception capacity of the deaf by discussions in real-life situations.

There will be a variety of applications where advanced processing of natural language is combined with speech synthesis. One of the new ideas is to use speech synthesis, morphological, syntactic and simple semantic processing to convert symbol language expressions (e.g. Bliss) into spoken messages of natural language.

Since the very beginning of the SYNTE 2 project at Tampere a reading machine for the blind has been one of the long term goals. A prototype of single-font text reader based on TV-camera and minicomputer is developed. This research system has still too low level of performance to be useful for practical purposes.

CONCLUSION

The devices and systems described here are examples of advanced communication, learning, or rehabilitation aids for the handicapped. The existing applications of SYNTE 2 have shown the usefulness of high-quality speech synthesis for these purposes. Technical progress through larger scale integration of speech processing hardware will lead to one of the most essential improvements, to acceptably low prices. Higher performance microprocessors and software will add more intelligence to the systems and devices. The present work at Tampere is directed towards a new, more language-independent, high-quality synthesizer and to new applications for it.

REFERENCES

(1) M.A. Karjalainen, "An Approach to Hierarchical Information Processes with an Application to Speech Synthesis by Rule", Doctoral Thesis, Acta Polytech. Scandinav., Ma 29, Helsinki 1978.

(2) M.A. Karjalainen, U.K. Laine, "Development of Communication Aids for the Handicapped Based on Speech Synthesis Techniques", Digest of 11th Int. Conf. on Med. and Biol. Eng., Ottawa 1976.

(3) T. Rahko, M.A. Karjalainen, U.K. Laine, and S. Lavonen, "Speech Audiometry by a Speech Synthesizer", Arch. Othorhinolaryngol. 222, pp. 85-89 (1979).

(4) M.A. Karjalainen, U. Laine, "Speech Synthesis Project in Tampere: Results and Applications", Proc. IV Nordic Meeting on Med. and Biol. Eng., Lyngby/Copenhagen, 1977.