Speech synthesisers for Linux,
Linux applications, blind users

Piotr PORWIK[*], Karol BĘDKOWSKI[**],
Łukasz ŻELECHOWSKI[*], Agnieszka LISOWSKA[**]

SPECIALISED LINUX SOFTWARE AIDING WORK OF BLIND PERSONS

For many years Linux environment has been evaluated from the only server system to desktop computer system. Thanks to this Linux has became serious competition for systems from the very popular Microsoft Windows family. Unfortunately sightless users of computer do not have wide choice of software speech synthesisers that may help them during working on computer in the Linux environment. The "Speakup" application is the one of the better ones designed for Linux. But, nevertheless, it does not support the maintenance in the same degree as popular applications for the Microsoft Windows. So in this paper the significant improvement of the "Speakup" is presented. This improvement allows blind users to quite comfortable work with the Linux system in the console text mode with the maintenance of Polish fonts, too. The application has become the user-friendly one with wide spectrum of possibilities, what many blind users confirm. The one of them is the co-author of this paper.

1. INTRODUCTION

Open operating systems made on the basis of Linux kernel gain bigger and bigger popularity, [1,2,3,4] but the availability of appropriate tools that enable blinds to use these systems is still quite insufficient. Unfortunately, creation of tools that enable blinds to work on with computer was interrupted due to lack of interest of commercial companies. The one of the better applications that aids work of blinds, which uses Linux as a main operating system, is the "Speakup". It operates the bash shell in the "screen echo" mode. In this mode the screen is read without content evaluation and also without analysis of displayed information context. Of course, the usefulness of such solution is limited. This project has not been perfected since 2001.

In this paper some modifications of the "Speakup" application are presented. Thanks to these modifications, Linux system tool and configuration programs as well as text editors from any distributions do not make barriers for blinds.

The main goal of the work described in this paper was to increase the range of available functions and to realise certain intelligence of the application in data reading. In the opinion of blinds, the maintenance of program packages is more crucial then the quality of generated speech.

[*] Institute of Informatics, Silesian University, Będzińska 39, 41-200 Sosnowiec, Poland
[**] Institute of Mathematics, Silesian University, Bankowa 14, 40-007 Katowice, Poland

Therefore the improved version of ”Speakup” uses available tools and programs generating synthetic speech not interfering in the sound quality.

The manner of modification of “Speakup”, described in the paper, is the work of two persons, the one of which is blind. Proposed software modifications follow immediately from blind community (working in informatics) needs. Unfortunately, non-blinds often not properly evaluate needs of this community. Therefore proposed changes may be not comprehensible or even needles for non-blinds.

2. MODERN SOFTWARE AND HARDWARE FOR BLINDS

From a point of view the Microsoft Windows systems family users Linux is the operating system that differs in architecture, application and method of usage. The operating system is the important part of each computer system. The last may be divided into the following parts: hardware, operating system and software. The hardware (processor, memory and I/O devices) makes up main resources of a computer system. The software (compilers, database systems, etc.) characterises the ways of usage of these resources to solve tasks made by users.

Operating system is the program that operates as mediator between computer user (together with its programs) and hardware. Operating system supervises and coordinates hardware maintenance by different application programs that work on order of different users. It also makes different operations on behalf of these processes.

The communication of blind user with computer is the versatile process. In simplification, this process may be divided into two stages (see Fig. 1):

- Analysis of screen’s content and state of keyboard (key pressed reaction),
- Transmission of processed data to device or program that can present it in a way accessible for blinds.

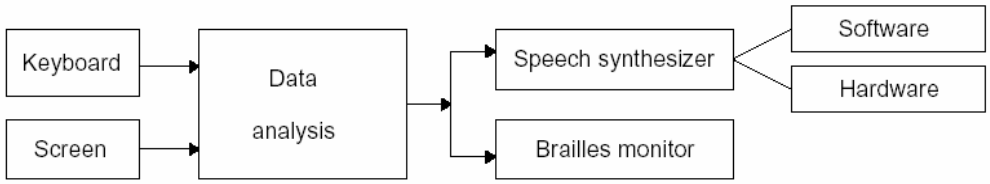

Fig.1. The process of communication between blind user and computer.

The data analysis is made by specialised software – the so-called “screen reader”. The transmission of information to blind takes place through generation of the synthetic speech with the help of speech generator (software or hardware) or through recording data from the “screen reader” with the usage of the Braille monitor. The analysis of changes appearing on the screen and the state of keyboard monitoring are not easy – not all information that appears on the screen should be read. Every time user should hear only that information, which needs his reaction at the moment and which should be controlled up-to-date. So the analysis of data appearing on the screen of the “screen reader” must characterise certain intelligence [5,9].

Nowadays the best blind user-friendly operating systems are the ones from the Microsoft Windows family. In particular, these are the functions of “Microsoft Active Accessibility” (MSAA)

that standardise the procedures of interpretation of individual objects on the screen. They make most functions of operating system available for blinds and offer maintenance of popular applications. The best ones are: “Jaws for Windows” from American Freedomscientific Company (*http://www.freedomscientific.com*) and “Window'eyes” from Gwmicro Company (*http://www.gwmicro.com*). Such solutions allow cooperating with many hardware synthesisers that is “Apollo II”, “Soundingboard”, and “Dectalk”. Additionally there is the possibility of installation of software speech synthesisers (from the Eloquence Company) that generate speech in eight languages, as well as software synthesisers based on the MSSApi technology. Moreover, the following Braille monitors may be used: “Alva Braille Terminal”, “Window Combi Braille” and “Braille Voyager” (*http://www.tiemangroup.com*).

Operating systems based on Linux kernel offer significantly fewer possibilities as regards adaptation of blinds needs. The companies delivering specialised hardware and software often ignore this problem. Also the appropriate program mechanisms needed for blinds are not included into the software. As a matter of fact many applications that help blinds to work with Linux have been implemented. Unfortunately, these programs often make up only enhancements of concrete applications and their usefulness is insignificant. The most popular enhancements are “emacsspeak” assisting “emacs” environment and “lynxspeak” that adapts “lynx” explorer to blind users.

The “Speakup” application has became the breakthrough. Thanks to it bash shell, commands may be accessible as sounds of synthetic speech in the “screen echo” mode. But this application reads only the screen’s content without any reaction on events. Such manner of working may be useful during usage of the basic shell commands, but is completely useless in tool programs, in which user needs specific functional fragments of the screen, for example menu items, file lists, the line above active cursor, etc. The application cooperates only with a few hardware speech synthesisers.

Another application that enables blinds to work in Linux is “brltty” (*http://mielke.cc/brltty*). This application operates most Braille monitors, including also the newest “Braille Voyager”. It runs as a separate service. Thanks to build in many Braille tables, it assures proper displaying of Braille’s alphabet in most popular languages. Its last version cooperates with such software synthesisers as “Festival” and “IBM ViaVoice”. But the main disadvantage of its usage is illogical manner of maintenance.

Hardware synthesisers often make up separate devices or enhancement cards. In some cases they are more comfortable in usage then software synthesisers, because not all computers have soundcard. Already dozen or so years ago hardware synthesisers were the only method of speech generation – computers had too small computational power to generate speech in real time. Nowadays such synthesisers are more rare in use because of their high unit price. The main advantage of such solution is large universalism and independence from system platform.

The most often-used hardware synthesiser on the Polish market is the “Apollo II” from Dolphin British Company (*http://www.dolphinuk.co.uk*). The device generates nice speech with slightly English accent and can speak in seven different languages. The interpreter of each language is placed in special ROM memory, installed into synthesiser as enhancement card. The synthesiser communicates with the computer through the serial port.

Nowadays the software synthesisers are the most popular methods of synthetic speech generation. The increasing computational power of computers allows applying algorithms that generate speech with much better quality, and then the one offered by hardware speech synthesisers.

Because software synthesisers need only a soundcard, such solution is much cheaper then hardware one. But the main disadvantage of that solution is dependence from particular system platform.

The most often used software synthesiser, designed for systems from Microsoft Windows family, is the "speak" synthesiser from Altix Company (*http://www.altix.pl*). It works in "Microsoft Text-To-Speech Engine" technology that ensures wide spectrum of configuration from speed of spoken word change through tone pitch change to control of interval between spoken words. But generated speech is metallic and hard to understand. Another software speech synthesiser that works under Windows system is "Gant" [7]. "Gant" generates nice speech in Polish and allows reading clipboard's content as well as text inserted to editing window. The application offers wide spectrum of parameters configuration and, as the only one from such applications, does not need installation process – it can be run directly from CD. The main disadvantage of it is lacking of possibility of integration the synthesiser with applications such as "screen reader" for Windows.

It has been made many speech synthesisers for the Linux systems family. The most of them is spread out with the GNU GPL license. In majority, these are software synthesisers for English speech and their usage for Polish speech is limited.

The best freeware synthesiser as regards speech quality is the "Festival" (*http://www.cstr.ed.ac.uk/projects/festival.html*). It makes available three language options: American, English and Spanish. Other languages, including Polish, are in the midpoint of implementation. The main advantage of this application is easy software side maintenance thanks to the build in lisp shell. Because of its wide possibilities and popularity most "screen readers" made for Linux environment use "Festival".

Another speech synthesiser, made for Linux, is "Say" (*http://www.inter.nl.net/users/jlemmens*). The application generates English speech, but thanks to patch named "orto2plf" – as only one from software synthesisers working under Linux control – it may speak in Polish. The quality of generated speech – in comparison, for example, with "Festival" – is poor. Unfortunately "say" is spread out without any documentation. Probably the application was made in February 1994 (such information is contained in the execution file).

Another software speech synthesiser, projected at the Prague University, is "The Epos Speech System" (*http://epos.ure.cas.cz/*). The application generates good quality speech in Czech and Slovak. User can choose in wide spectrum of voices and thanks to configuration files he also may extend the application with his own dictionaries.

Nowadays, the most advanced, software speech synthesiser is "RealSpeak"" from ScanSoft Company (*http://www.scansoft.com*). The application generates high quality women's voice speech in one of nineteen languages (including Polish). The application operates on Windows platform and, from recently, also in Linux systems. Because of the high price (over 1000 Euro) it is used mainly in telecommunication technologies (e.g. reading short messages in wireless telephony).

3. "SPEAKUP" AND ITS ADAPTATION TO WORKING WITH BLINDS

Only few computer applications are adapted to cooperation with blinds. The ignorance of needs of these people is unacceptable nowadays. The one of important applications from Linux environment is the "Speakup". The application, communicating directly with the shell, aids blind's work by reading communicates displayed on the screen, as well as fragments of text pointed by

cursor. Unfortunately, this program, in its distributive version, is not well adapted to working with blind users. The conveniences offered by this program are insufficient, what was confirmed by many blind users. So the aim of the work presented in the paper was such modification of the "Speakup" application, that it becomes user friendly for blind programmers working in Linux environment.

"Speakup" is the one of the "screen reader" type applications for Linux platform. It analysis the screen's content and keyboard's state, and basing on it, controls the speech synthesiser. The connection of "Speakup" with other elements of computer system is shown in Fig. 2.

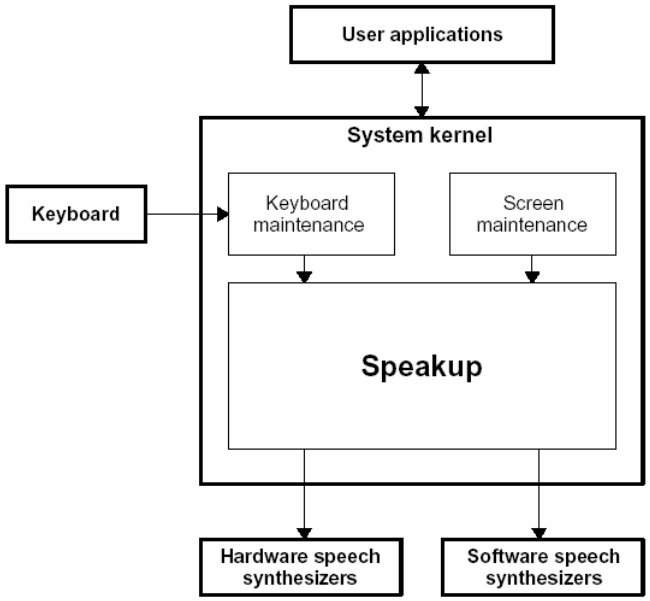

Fig.2. Location of "Speakup" in the process of communication with blind user.

"Speakup" is probably the only one application of the "screen reader" type ones designed for Linux systems, in which from the beginning the main goal has been to assure universality and independence from run programs.

The application was modified by adding new functions extending its possibilities. The principles of screen's content analysis were changed also. Of course, these modifications do not violate the spread out principles based on GNU GPL license.

In the distributive version "Speakup" have the following possibilities:

- It works in "screen echo" mode – reads all information directed by applications to screen – this mode is useful only during direct working with shell;
- It works in navigation mode – reads fragment of screen pointed by user;
- It processes only ASCII codes (code page 437) – national fonts are ignored;
- It cooperates only with hardware speech synthesisers;
- In the case of using window interface applications it reads some fragment of screen;
- The configuration of application and speech synthesiser may be performed with the help of virtual "proc" file system.

Performed improvements of the application allow using the following new functions in "Speakup":

- Maintenance of “ncurses” library based applications, by intelligent reading lighted menu position as well as reading screen’s content;
- Automatically reading the name of chosen file in “Midnight Commander”;
- Maintenance of any text editor;
- Cooperation with software speech synthesisers, by adding virtual device “say”. This solution allows to use any software synthesiser available for Linux platform;
- Adding “speakup_festival” and “speakup_say” allows to cooperate with software speech synthesisers “Festival” and “Say” respectively;
- Adaptation of “Speakup” to reading text with all Polish fonts;
- Adaptation of hardware speech synthesiser driver “Apollo II” to reading Polish fonts;
- Modification of keyboard mapping of “Speakup” that allows to write Polish fonts;
- Adding “speakup_config” has simplified program configuration and ensures keeping all settings to be automatically restored any time;
- Adding “speakup_man” has made easier the usage of system documentation (man pages) for blinds.

Also much less significant changes were made to improve the functionality of the application. Users, especially the blind ones, intensively tested all mentioned functions. The choice of functions was determined by the opinions of blinds.

4. THE PRINCIPLES OF WORKING OF THE MODIFIED “SPEAKUP” APPLICATION

In the installation process “Speakup” combines with the system kernel and becomes its integral part. It causes that it is active directly after starting the system kernel, so blind user may get information generated during starting of individual system components. The application ends its work along with system kernel and informs user about events generated during shutting down the system. The scheme of “Speakup” is shown in Fig. 3.

Within the boundaries of the work described in the paper to “Speakup” was added the module of software synthesiser driver with the virtual “say” device connected to it. Moreover the module responsible for data analysis was modified. The new module is also the one that operates Polish fonts as well as their conversion on appropriate standard needed by speech synthesiser.

The application catches commands coming from application programs to kernel, which concern changes of screen’s content and information about key pressed. This information is then further processed and sent to currently chosen speech synthesiser driver. The driver adapts the data to needs of synthesiser and sends them to peripheral device (speech synthesiser).

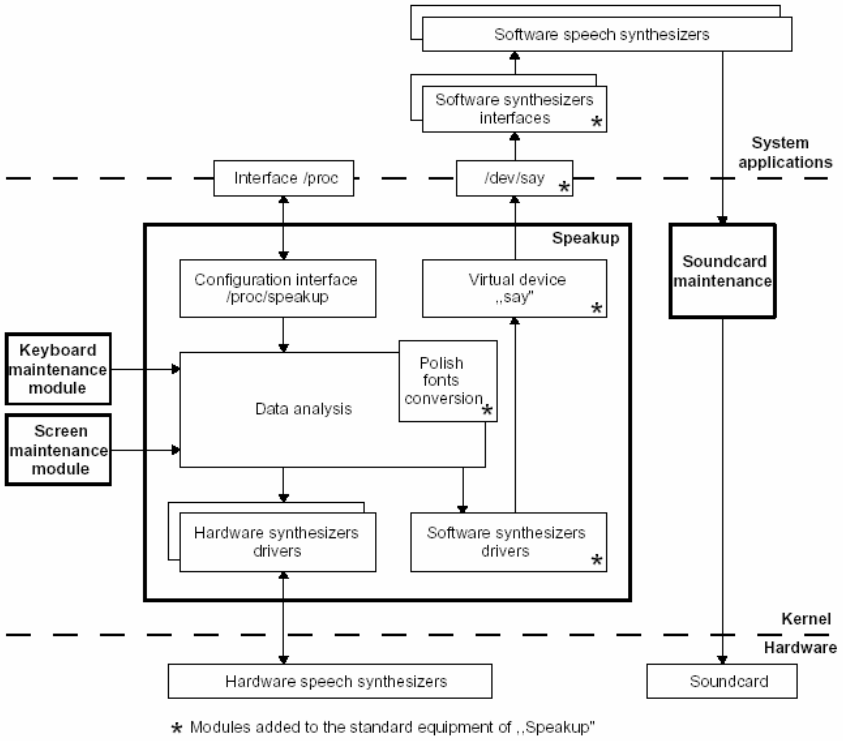

Fig.3. The principle of modified "Speakup" working.

The driver "say" and the dedicated virtual device "say" are the event. They deliver data for software speech synthesisers. Data from "Speakup" contain also (beside the text to reading) control sequences. These sequences must be appropriately filtered. Additionally different types of software synthesisers need different manner of getting the data. So it is necessary to use additional applications that make up connection between the driver and the synthesiser. In the "speakup_utils" package (the kit of our specialised programs) it were included two programs: "speakup_say" for "Say" synthesiser and "speakup_festival" for "Festival" synthesiser.

The modified application of "Speakup" may work in four modes:

- "Screen echo" mode – reads all fonts appearing on the screen;
- Navigation mode – it allows the user to introduce the screen's content at once. Thanks to specialised virtual cursor this process does not disturb working applications. This mode offers some possibilities: reading the line above, below and beneath virtual cursor, reading word on the left, right and below the cursor, reading words spelling out, reading data from upper part of screen to virtual cursor position, reading from beginning of the line to the cursor position and reading from cursor position to the end of line;
- Cursor tracking – it allows to speak fonts nearby the active cursor. It is especially used during working with text editor (all editorial operations are made with usage of the active cursor);
- "Midnight Commander" mode that allows to use the "Midnight Commander" application ("mc") and, in particular, the menu items included in "mc", as well as the text editor included in "mc" ("mcedit"). This mode allows working with other applications based on "ncurses" library too.

The modifications of the sources of “Speakup” are not sufficient to assure the appropriate functionality of the application. Not all parts may be included in the kernel, due to its size limit. That is why a part of functions has been placed outside the kernel. The connection between “speakup_utils” package and “Speakup” is shown in Fig. 4.

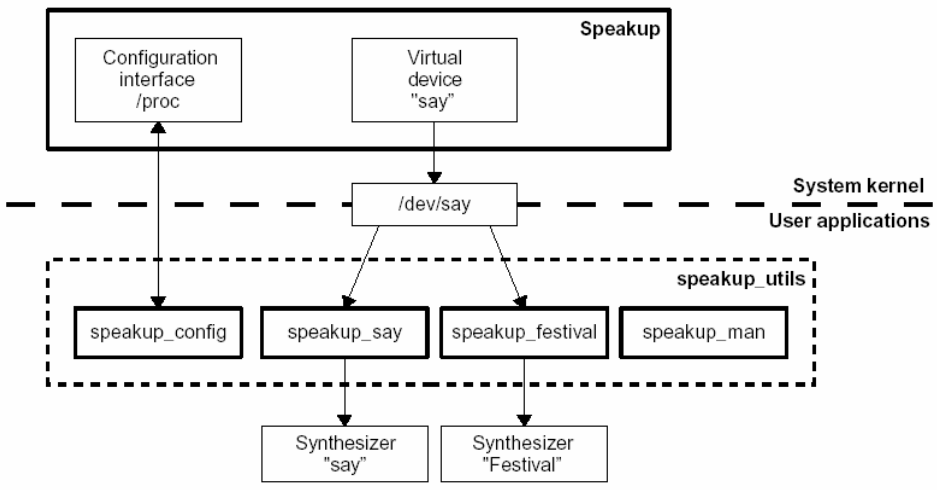

Fig.4. The connection of modified application “speakup_utils” with other applications.

5. THE ADVANTAGES AND DISADVANTAGES OF THE MODIFIED PROGRAM “SPEAKUP”

The modified version of “Speakup” has the following advantages:

- The integration with the system kernel and usage of hardware speech synthesiser ensure the contact of blind user with system beginning from starting system till shutting down;
- Small hardware requirements;
- Maintenance of any text shell;
- Maintenance of software synthesisers (“Festival” and “Say”) and popular hardware synthesisers;
- Maintenance of “mc” by reading menu items and reading names of chosen files or directories;
- Maintenance of any text editor;
- Making available man pages for blind users;
- Reading screens in navigation mode.

And the main disadvantages of the improved application are as follows:

- Difficult installation process, it demands from the user good knowledge of the Linux environment;
- The application works only in text mode of the system.

Among the systems from Linux family the only one application that may compete with the “Speakup” is “brltty”. This application operates wide spectrum of Braille monitors. Moreover it runs as separate daemon – it releases user from the kernel compilation process. The lack of

integration with kernel causes that blind does not have contact with the system during system starting process. Its maintenance is very complicated and in a large degree dependent on possessed Braille monitor.

"Speakup", as the only one from the "screen readers" for Linux systems, holds easy to use module that makes possible to remember user's preferences. Its maintenance depends on choosing individual functions from menu, however "brltty" requires the knowledge of many key combinations, which are established by means of Braille monitor.

BIBLIOGRAPHY

[1] BOWMAN I. T., HOLT R. C., BREWSTER N. V., Linux as a Case Study: Its Extracted Software Architecture, Proceedings of ICSE-99, Los Angeles, 1999.
[2] DE LA RUE M., Linux Accessibility HOWTO, *http://www.tldp.org*.
[3] Linux Kernel Hacking HOWTO, *http://netfilter.kernelnotes.org/kernel-hhacking*.
[4] MC CARTY B., Debian GNU/Linux, Helion, Gliwice, 2002.
[5] O'MALLEY M. H., Text-to-speech Conversion Technology, Computer, August, pp. 17-23, 1990.
[6] POMERANTZ O., Linux Kernel Module Programming Guide, *http://www.tldp.org/LDP/lkmpg*.
[7] PORWIK P., SZCZEPANKIEWICZ M., The Voice Synthesiser of Polish Text to Blind Persons, Journal of Medical Informatics and Technologies, Vol. 4, pp. MT-101-MT-109, 2002.
[8] RUSLING D. A., The Linux Kernel, *http://www.linux.doc.org/LDP/tlk/tlk.html*.
[9] WITTEN I. H., Principles of Computer Speech, Academic Press, 1992.