RAINBOW BRIDGE

Training Center based on Voice Technology for People with Physical Disabilities

Jan Nouza, Petr Cerva and Josef Chaloupka
Institute of Information Technology and Electronics, Technical University of Liberec, Studentska 2, Liberec, Czechia

Keywords: Voice technology, Motor disabled persons, Speech recognition, Hands-free control.

Abstract: The Rainbow Bridge is an EU-supported project that aims to teach Czech people with physical disabilities to benefit from voice technology products in their lives and prospective jobs. All products - voice control tools and dictation programs - have been developed in our lab during the last decade. At this stage, the main goal is to teach the target group how to use PCs in hands-free manner (though some have no previous experience with computers), how to adapt the tools to their specific speech abilities and how to acquire the skills that can be useful for their potential employers. One of the lecturers is a quadriplegic person who herself has been using our voice products for almost 5 years.

1 INTRODUCTION

Voice technology has a great potential as an alternative means for e-inclusion of people with physical disabilities. This was demonstrated already in the early years of computer speech processing (Noyes et al., 1989) and it is even more evident now when operation systems with embedded speech input and output have become widely available. At the end of the first decade of the 21st century, voice is considered as an optional and complementary means in human computer interaction.

However, there is a big difference between common PC users and those with physical disabilities (Hawley et al, 2005). For people whose impairment has not allowed them to use computers so far, it is not easy to start now and learn by themselves. Recent SW applications utilize sophisticated (mostly graphically oriented) user interfaces where the mouse and keyboard are the primary tools for interaction. When voice commands are available, they often mimic the actions of these standard tools. This is natural for the majority of the users but difficult to understand for the people who have never worked with computers before because their impairment had not allowed it. Should they learn the computer basics in the same way as the other users? Or should we provide them with special courses where the alternative interaction tools eligible for them, such as voice input and output, would be presented as the primary ones? These and several other questions were raised when we were asked to help in preparing a project of a training center aimed at physically disabled people who want to start with computers.

The need for such a training center has become urgent when voice recognition programs designed for Czech motor-handicapped people appeared on the market. In 2005 it was MyVoice - a voice-controlled interface for PCs (Nouza et al, 2005), in 2007 MyDictate - a very-large vocabulary discrete dictation tool (Cerva et al, 2005), and in 2008, NewtonDictate - a program for fluent dictation with 300K+ word lexicon, all developed in our research team. Only a small part of the potential users have been able to learn to use them effectively. It was mainly the people who had already worked with computers before they suffered some sort of motor impairment. Some others succeeded because they found helpful assistants within their families or friends. But many others just gave up after a short initial trial that did not bring immediate success due to various reasons, such as disordered speech, wrong usage or little patience. And last but not least, many potential users have not learned about the tools, yet.

In 2009, the project called Rainbow Bridge was prepared and submitted to the EU-supported Operational Program Human Resources and Employment. Its goal was to build the training center in Prague where people from the target group could learn about the available technology, try it under the supervision of professional teachers,

Nouza J., Cerva P. and Chaloupka J..
RAINBOW BRIDGE - Training Center based on Voice Technology for People with Physical Disabilities.
DOI: 10.5220/0003147905290533
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2011), pages 529-533
ISBN: 978-989-8425-34-8

receive individual care and assistance, and also share their experience with the others. One of the top priorities was to give at least some of the trainees a chance to get a job where they could employ their new skills. Eventually, the project was approved and the center launched its activities in 2010.

In the following sections, we first describe the set of tools developed for the clients of the center, then we present the teaching methods and finally we discuss the experience gained during the preparation of the courses.

2 VOICE CONTROLLED TOOLS

Here we briefly present the individual tools that have been developed for and are being used in the training center.

2.1 Voice Interface for PC

Program MyVoice (Nouza et al, 2005) serves as an interface between commands spoken to the microphone and standard applications running on a PC. It either complements or (in most cases) replaces keyboard and mouse as shown in Fig.1.

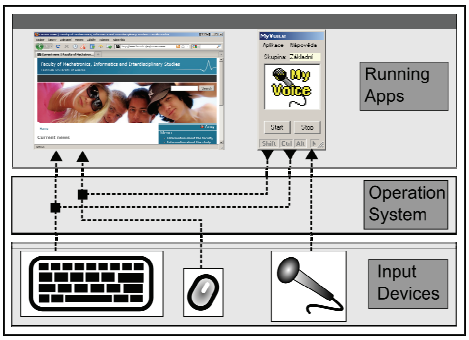

Figure 1: Voice control of PC applications - basic scheme.

The Myvoice is mostly used as a tool that launches and controls applications, enables typing (of letters, frequent words and phrases), and provides access to internet services. Anyway, it allows for virtually any activity that is normally performed via a keyboard and mouse. Compared to the first release in 2005, the latest version includes some new features. Let us mention at least the improved navigation on web pages, either by adding numbers to the web links or by the introduction of a target-like grid that simplifies mouse cursor movements. It is also possible to hide MyVoice's window, which may be useful, for example, when watching video in the full-screen mode.

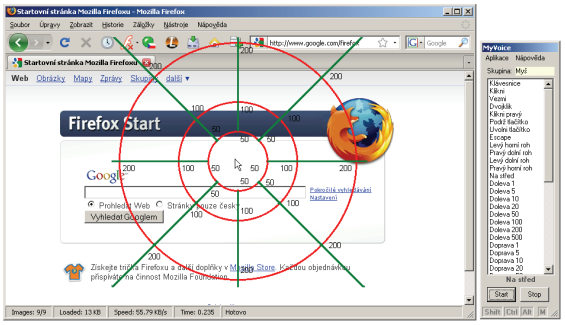

Figure 2: A snapshot of voice controlled internet browser.

2.2 Voice Interface for Home Devices

The MyVoice tool can be further expanded to allow for controlling common home devices, like e.g. lights, electrical heating or air conditioning, window curtains, doors, or audio and video equipment.

The extension module includes a hardware unit that contains two types of transmitters: one whose signal is compatible with wireless relays used to power home electric appliances, and an IR transmitter to control a TV set, stereo, video and a satellite receiver. The HW unit is linked to a PC via USB connection. Not only the devices, but also the user can communicate with the control program in a wireless way. Our tests showed that even the cheap Bluetooth microphones, either headworn or lapel ones, proved to work well with the MyVoice.

In 2007 we built a demonstration room where all common home devices were controlled via the MyVoice and wireless links (Nouza et al, 2007). Its diagram is shown in Fig. 3. The same arrangement will is used in the training centre in Prague.

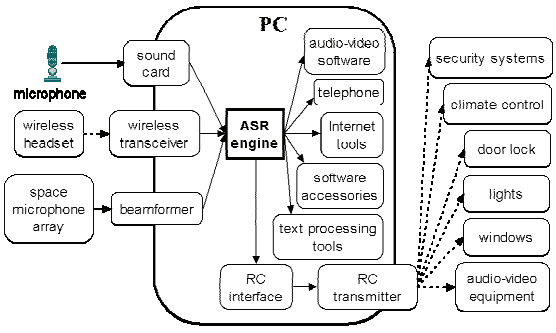

Figure 3: Voice control extended to home devices.

2.3 Hands-free Dictation Programs

During the last 5 years we have developed two types of dictation programs, one for discrete speech input, another for natural fluent speech.

2.3.1 Discrete Dictation

The program for discrete speech input is called MyDictate. Its design has been fully determined by the target group, i.e. by the people who cannot use their hands. Therefore, not only the dictation itself but also all supporting activities, like editing, error correction, formatting, or lexicon maintenance had to be designed as hands-free operations. The program has been described in detail in (Cerva et al, 2005).

The SW is distributed with a general purpose lexicon containing about 550 thousand words. The employed technology allows the program to run even on recent low-cost PCs. When a word is uttered, the recognizer outputs the ordered list of 10-best candidates, taking into account the acoustic as well as the language model score. The word with the best score is automatically added to the dictated text while the next 9 candidates appear on the list shown in MyDictate's window. In case of lexical ambiguity or when a minor recognition error occurs, the user can take another candidate from this list and replace the wrongly typed item. There are about 100 other control commands that can be used e.g., to delete the last character(s), word(s), or sentence(s), to select a part of text, to work with the clipboard, to move the cursor, to spell the individual letters or to toggle their (lower or upper) case. The basic vocabulary assures about 99 % coverage rate for common Czech texts. If a dictated word is not in the lexicon, the user can add it by voice during the dictation session.

As the program is aimed particularly at people with physical disabilities, it must be able to cope with less standard pronunciation. If this is the case, the user can employ the embedded speaker adaptation module. It prompts him/her to say 300 phonetically balanced words. For most users it helps to reduce the recognition error up to 25 % relatively.

2.3.2 Fluent Dictation

The software for fluent dictation was developed by in collaboration with the Newton Technology company, which distributes it under name NewtonDictate. This software is aimed at general public, and at professions, like lawyers, doctors and people from the media domain. It comes with several types of lexicons. The general-purpose one is the largest and it contains 500K words recently. The profession oriented lexicons are smaller (320K for lawyers and 140K for medicine) with domain specific language models.

Though originally, this program has not been designed for hands-free use, it has been considered for the lectures and training sessions in the center. One reason is that some of the clients showed their interest in learning this software and exploiting it in their prospective jobs. It was mainly those people whose physical disability does not affect their speech and who can use their hands at least to some extent. For the other clients, integration with the MyVoice is being prepared. It is truth, however, that many persons with disabled hands prefer the discrete dictation to the fluent one. They appreciate namely the facts that a) they can form the text in their own pace (while the fluent speech technology requires more or less continuous flow of words), b) they can correct or modify the input text immediately, c) they have feeling that the isolated-word decoder is more robust to speaker-produced and background noises as well as to hesitation sounds, and last but not least d) they can easily add new words to the lexicon. It should be also noted that in Czech - because of its rich and complex morphology - many word-forms differ only in one or two characters, which means that they sound very similar and can be easily confused, particularly in fluent speech. For the users, correcting these small errors spread within a fluently input text is a frustrating task, if it must be done in hands-free manner. Similar observations were reported also in (Hawley et al, 2005).

3 TRAINING CENTER

The Rainbow Bridge project has been supported by a 200K Euro grant, from which one half has been spent by building the training center and the other will cover the running costs of a three-year pilot operation. The center is situated in Prague in place with good access by public and private transport.

3.1 Project Goals

The main goals of the project are:

Promotion of the voice technology among those people with physical disabilities who can use it as an alternative means of interaction with computers.

Creation of a pilot training center with certified teaching methods (which can be later replicated on regional levels).

Teaching the basic computer skills to the people whose disability had never allowed them working with PCs.

Giving at least some of the trainees a chance to employ PCs together with the voice technology in their prospective jobs, e.g. in call and help centers, in voice-scanning of documents, in re-speaking

tasks, etc.

3.2 Trainees

A national survey (CSO, 2008) estimates that the people with physical disabilities make up about 5 % of the total population in the Czech Republic. In the Prague region, this means circa 50,000 persons, from which about 5,000 suffer from complete or partial hand paralysis. This pilot project can offer a chance to 50 of them.

Because of this limited number, the selection of the course participants is done so that people with different levels of motor impairment, different computer skills and different voice ability can participate in multiple course runs. Allowing this rather wide spectrum of clients gives this pilot project a good opportunity to develop and verify various general and specific training methods.

3.3 Teaching Staff

The teaching and project supporting staff consists of several professions: a teacher specialized on education of disabled people, a teacher familiar with ICT and voice technology, a psychologist, a voice therapist, and a lecturer who himself or herself uses the voice technology tools in his/her daily living.

During the first year, the latter position was offered to a quadriplegic woman who has been using the MyVoice and MyDictate software since their first dates of release. She has utilized them for her high school study (organized from her home and successfully completed in 2009). Recently, she has become known also for her blogs published regularly in a major Czech internet newspaper.

3.4 Equipment

The center consists of one room with 8 work places, each equipped with a PC and voice technology software. The room is used for two types of activities: First, for the lectures when up to 8 trainees can participate and watch the presentations and demonstration on their monitors, and second, for training sessions when no more than 2 students work on assignments under the supervision of a teacher. The center also owns 8 laptops with the same software equipment. These can be lent to the clients for their home work. The room is equipped with the remote controlled devices mentioned in section 2.2.

4 TRAINING METHODS

4.1 Pilot Phase of Training Courses

The project started in January 2010 and it will end in December 2012. Within the first six months, a pilot phase with 2 trainees is running. One of them is a person who suffered a serious spinal injury two years ago but before that he had worked with a PC. The other person also cannot use her hands but, unfortunately, she has not got any prior knowledge of computers. Moreover, she suffers from moderate dysarthria.

These two people were the subjects in the preparatory phase of the courses. The fist trainee needed very short time to learn how to replace the keyboard and mouse by voice in the activities he had been doing before his injury. After a 4-week training he has been able to use the MyVoice and MyDictate programs very efficiently. Currently, he learns to exploit the fluent dictation tool as well. His experience and feedback helps us in completing the integration of the dictation and voice command control tools.

The other person requires much more individual care. In her case, the MyVoice tool has been tailored and adapted to her specific speech. The command groups have been modified in the way that they included significantly smaller numbers of task oriented commands, like basic cursor movements, easily pronounceable and distinguishable names of Czech letters, commands for basic navigation on web pages, etc.

4.2 Standard Courses

Standard courses start in September 2010. Each course run has 8 participants and lasts for 3 months. It consists of 1 month period of lecturing and basic training (60 contact hours in the center + 40 hours homework) and 2 months of individual practicing (80 contact + 80 homework hours). The practical sessions are organized so that one supervisor is available for two trainees.

The curriculum includes these main topics:

- introduction into PCs, navigation within the MS Windows platform (alternative use of desktop, Start button, shortcuts, mouse actions, window actions, etc.),
- using the MyVoice instead of a keyboard and mouse, hints for correct and efficient use of voice, optional adaptation to speaker's voice, practicing and developing the voice-command skills on simple applications, such as the Solitair,

- game, the Calculator, the Paint tool, a music player, etc.
- text typing and editing within the MS Word program using the tools for discrete and fluent dictation,
- internet and web services, using the tools for web browsing, e-mail and IP telephony communication, social networks.

The topics are introduced by a teacher first and then presented by the assistant lecturer (usually one of the previously trained handicapped users) who explains the practical issues of the voice-command actions using his or her own previously acquired experience.

4.3 Course Benefits

Each of the course participants will learn how to exploit voice technology tools for his or her daily activities, for study, for leisure, for establishing internet contacts with other people, or for emergency situations. Moreover, we believe that at least some of the graduates will be given a chance to use their new skills in suitable jobs. The center itself will offer a temporary lecturer job to the best course participant. There are also provisional agreements with several institutions in Prague where the best graduates can be employed, e.g. as assistants in call and help centers or as re-speakers in companies that specialize on transcription of broadcast or meeting speech.

5 CONCLUSIONS

This project is the first activity of this type organized on the national level and supported by the European Union. It gives a new chance to the people who have not been able to use modern information and communication technology so far because of their physical handicap. It is also a great opportunity for the researchers in the speech technology domain to see how their products perform in rather challenging conditions and how these products can be further improved. If the project succeeds, similar centers will be built also on regional levels.

ACKNOWLEDGEMENTS

The research reported in this paper was partly supported by the Czech Science Foundation in project no. 102/08/0707 and by the EU-supported CZ.2.17/2.1.00/32644 grant in programme Prague – Adaptibility.

REFERENCES

Noyes, J., Haigh, R., Starr A., 1989: Automatic speech recognition for disabled people. *Applied Ergonomics, Vol.20, no.4, pp. 293-298.*

Hawley, M.S., Green, P., Enderby, P, Cunningham, S., Moore, R., 2005: Speech technology for e-inclusion of people with physical disabilities and disordered speech. *Proc. of Interspeech 2005, Lisboa, pp.445-448*

Nouza, J., Nouza, T., Cerva, P., A Multi-Functional Voice-Control Aid for Disabled Persons". *Proc. of Specom 2005, Patras, 2005, pp.715-718.*

Cerva, P., Nouza J.: Design and Development of Voice Controlled Aids for Motor-Handicapped Persons. *Proc. of Interspeech, Antwerp, 2007, pp. 2521-2524.*

Nouza, J., Chaloupka, J., Zdansky, J., Silovsky, J., Kroul, M., Mader, Z.: Voice Controlled Center for Homes of Motor-Handicapped Persons. *Proc. of Specom 2007, Moscow, 2007, pp. 714-719.*

Desilets, A.: VoiceGrip: A Tool for Programming-by-Voice. *International Journal of Speech Technology. Vol. 4, No. 2, June 2001, pp. 103-116.*

CSO (Czech Statistical Office): Results from survey of disabled persons in 2007. *Published by Czech Statistical Office in 2008 and available on web: http:/www.czso.cz/csu/2008edicniplan.nsf/t/4100269DD7/$File/330908j3.pdf*