EVALUATION OF AUDITORY REPRESENTATIONS FOR SELECTED APPLICATIONS OF A GRAPHICAL USER INTERFACE

György Wersényi

Széchenyi István University
Department of Telecommunications
H-9026, Győr, Hungary
wersenyi@sze.hu

ABSTRACT

A survey with 50 blind and 100 sighted users included a questionnaire about their user habits during everyday use of personal computers. Based on their answers, the most important functions and applications were selected and results were compared. Special user habits and needs of blind users are highlighted. The second part of the investigation included collecting of auditory representations (auditory icons, spearcons etc.), mapping with visual information and evaluation with the target groups. Furthermore, a new design method for auditory events and class was introduced, the so called auditory emoticons. These use non-verbal human voice samples to represent additional emotional content. Blind and sighted users evaluated different auditory representations for the selected events, including Hungarian and German spearcons. Finally, an application can be created and sound samples can be implemented under JAWS or the Windows OS.

1. INTRODUCTION

Screen readers and text-to-speech applications serve as the main applications for blind users. Many of them offer the possibility of setting the speed of speech, but non-speech sounds are seldom applied. However, earcons or auditory icons can be a powerful tool to represent visual content end events of the screen [1, 2]. Furthermore, spearcons have been proved to be a good alternative for menu navigation or replace simple "speeded up" speech [3, 4].

In the frame of the GUIB (Graphical User Interface for Blind Persons) project an auditory user interface was developed. In this case, the most important and frequently used icons, events and applications had to be determined, furthermore, possible sound representations had to be found as a replacement or extension to the visual content.

After determining the most important functions and applications, a collection of sound samples has been developed and evaluated based on the comments and suggestions of blind and sighted users. This set of auditory icons, earcons and spearcons is available in wave or mp3 format.

This paper presents a collection of sounds that have been selected by the users as the "winning" versions. Detailed results, evaluation rates and numbers are not shown. The sound samples are included or can be downloaded from the Internet.

Although, the primary target group is the blind community, sighted users can also benefit from auditory feedbacks of the operation system.

1.1. Auditory Representations

There are different possibilities for creating auditory representations. First, we can use auditory icons. These sounds are short „icon-like" sound events having a connection in meaning to the visual event they represent. Auditory icons are easy to interpret and easy to learn, therefore, we can use more of them even simultaneously. Users may connect and map the visual event with the sound events for the first time they hear them. Typical example is the sound of a matrix dot-printer that everybody connects somehow with printing (in progress). Unfortunately, there are events on a screen that are very hard to represent by auditory icons.

On the other hand, earcons are „meaningless" sounds. The mapping is not obvious, so they have to be learned together with the event they are linked to. E.g. the sounds that we hear during start-up and shut down the computer or during warnings of the operation system are well-known after we hear them several times. They are harder to interpret and to learn.

Spearcons have already proved to be useful in menu navigations and in mobile phones because they can be learned and used easier and faster than earcons [5]. Spearcons are speech samples specially compressed in time, often names, words or simple phrases. Compress ratio, required quality and spectral analysis was made for Hungarian and English language [6]. Besides the Hungarian samples, the German spearcon database for our study was created with native speakers (see later).

Based on the evaluation of the survey with the users, we will introduce a new group of auditory events called auditory emoticons. Emoticons are widely used in e-mails, chat and messenger programs, forum posts etc. The different smileys and abbreviations (such as brb, rotfl, imho) are used so often that users suggested to represent these with auditory events as well.

Furthermore, some sound samples can not be classified into the main three groups mentioned above. Auditory emoticons are non-speech human voice(s), sometimes extended and combined with other sounds in the background. They are related to the auditory icons the most, using human non-verbal voice samples with emotional load. Auditory emoticons – just like the visual emoticons - are language independent and they can be interpreted easily. E.g. the sound of laughter or crying can be used as an auditory emoticon.

All auditory events are meant both for sighted and blind users as feedback of a process or activation, to find a button, icon, menu item etc.

2. EVALUATION AND COMPARISON OF USER HABITS

In order to find the most important applications and functions, we prepared a detailed questionnaire both for blind persons as well as for users with normal vision for comparison. The measurement method and some preliminary results of sighted users were presented and described in [6]. Furthermore, user habits of different age groups and user routine were investigated. At the end, the most important features, applications were selected in order to create auditory representations for them as an extension or replacement for speech. This evaluation should deliver important information about applications and sub-functions, and furthermore, about differences between blind and sighted persons.

The survey included 100 persons with normal vision and 50 visually impaired (from Hungary and Germany). Subjects were categorized based on their user routine and based on their ages as well. 83% of the subjects were "average" or "above average" users for sighted people but only 40% is this for blind users. It is clear that blind users often restrict themselves for basic use.

Average age of sighted users is 27,35 years and 25,67 of blind persons. Subjects had to be at least 18 years of age and they had to have at least basic knowledge of computer user routine. During evaluation a simple mean average was calculated based on the given points. Ranking points are from 1 point (least significant) to 5 points (most significant). Average points above 3,75 correspond to frequent use. On the other hand, rates below 3 points on average are regarded not to be very important. Because some functions appear several times on the questionnaire, these rates were averaged again (e.g. if „print" has an average value of 3,95 in Word; but only 3,41 in the browser then a mean value of 3,68 will be used).

Summarized averaged results are listed in Table 1. Yellow marked fields indicate important and frequently used programs and applications (3,00 – 3,74). Light green fields indicate everyday use and higher importance (above 3,75 points). Similarly, left column's filling indicates whether the application is important for only sighted, only for blind or for all users. Dark green filling is used when one is light green and the other is yellow of the right columns, in any other cases the colour matches the other columns' filling.

At the end of the table some additional ideas and suggestions are listed without rating. Further suggestions from sighted users were: Wave editor, remove USB stick, date and time. Blind users mentioned DAISY (playback program for reading audio books), JAWS and other screen-readers. The frequent use of emoticons (smileys) in e-mails and messenger applications brought up the need to find auditory representations for these as well.

2.1. Blind Users

Blind users have different needs sometimes when using personal computers. We observed that

- Blind users like the icons, programs that are on the desktop by default, such as My Computer and the My Documents folder. They use these more frequently than sighted users, because sighted can easily access other folders and files deeper in the folder structure as well.
- Programs that use graphical interfaces (e.g. Windows Commander) to ease the access are only helpful for sighted users.

- Image handling, graphical programs, movie applications are only important for sighted users. However, the Windows Media Player is also used by the blind persons, first of all for music playback.
- Select and highlighting of text is very important for the blinds, because Text-To-Speech applications read highlighted areas.
- Blind users do not print often.
- Acrobat is not popular for blind persons, because screen-readers do not handle PDF files properly. Furthermore, lots of web pages are designed with graphical contents (JAVA applications) that are very hard to interpret.
- Word is important for both groups, but Excel, Power Point use mainly visual presentation methods, so these latter programs are useful for sighted users.
- For browsing the Internet, sighted users like more the "new tab" function, while blind persons prefer the "new window" option. It is hard to orientate for them under multiple tabs.
- The need for gaming was mentioned by the blinds as a possibility for entertainment (specially created audio games).

The idea of extensions or replacements of these applications by audio signals was welcomed by the blind users, however, they suggested not to use too much of them, because this could lead to confusion.

Blind users mentioned that JAWS and other screen readers do not offer changing the language "on the fly" so if it is reading in Hungarian, all the English words are pronunciated phonetically. This is very disturbing and makes the understanding difficult. Interesting is that JAWS 9.0 does not offer yet Hungarian language, so Hungarian blind users use the Finnish module. This is the best method to understand the speech, although, the reputed relationship between these languages has been refuted lately. Another complaint was that JAWS is expensive and the free version of a Linux-based screen reader has a low quality speech synthesizer.

The GUIB project tested Braille input modules, but nowadays they are not used anymore. Furthermore, the world of Braille is limited for blind users only, sighted have no access.

The best way for a blind person to access applications would be a maximum of a three-layer structure (in menu navigation), alt tags in pictures, and the use of the international W3C standards (World Wide Web Consortium) [7]. Only about 4% of the internet web pages follow these recommendations.

As blind users mentioned before, there is an extreme need for gaming and entertainment. This means audio-only games. Text-based adventure games using the command line for navigation and for actions are also very popular. But there is more need for access to on-line gaming, especially for on-line table and card games, such as Poker, Hearts, Spades or Bridge. This could be realized by speech modules, if the on-line website would tell the player the cards he holds and are on the table.

One of the most popular is the game *Shades of Doom*. A trial version can be downloaded from the internet [8]. In a three dimensional environment, you must guide your character through a research base and shut down the ill-fated experiment. It features realistic stereo sounds, challenging puzzles and action sequences, original music, on-line help, one-key commands, five difficulty levels, eight completely navigable and explorable levels, the ability to create Braille-ready maps and much more. This game is designed to be completely accessible to blind and visually impaired users, but is compatible with JAWS and Window Eyes if desired.

Programs/applications/ functions	sighted	blind
Number of subjects	100	50
	Total avg.	Total avg.
Internet Browser (icon/starting of the program)	4,62	4,67
E-mail client	4,11	4,67
Windows Explorer	3,98	3,25
My Computer	3,58	3,94
Windows/Total Commander	2,75	1,88
Acrobat (Reader)	4,26	3,33
Recycle Bin	4,41	3,67
Word (Word processor)	4,53	4,56
Excel	3,81	2,94
Power Point	3,14	2,47
Notepad / WordPad	2,6	2,61
FrontPage (HTML Editor)	2,09	2,24
CD/DVD Burning	3,84	3,94
Music/Movie Player	4,09	4,17
Compressors (RAR, ZIP etc.)	3,41	3,22
Command Line, Command Prompt	2,84	1,83
Printer handling and preferences	3,87	2,64
Image Viewer	3,98	1,62
Downloads (Torrent clients, DC++, GetRight)	2,7	2,06
Virus/Spam Filters	4,29	4,39
MSN/Windows Messenger	3,29	3,17
Skype	2,91	3,39
ICQ	2,33	1,78
Chat	2,06	1,5
Paint	3,05	1,67
Calculator	2,82	2,61
System Pref., Control Panel	3,6	3,17
Help(s)	2,55	2,98
Search for files or folders (under Windows)	3,32	3,11
My Documents folder on the Desktop	3,52	4,11
JAWS, Screen-readers	1	4,6
FUNCTIONS		
Home button (Browser)	3,53	3,61
Arrow back (Browser, My Computer)	4,22	4,33
Arrow forward (Browser, My Computer)	4,22	4,33
Arrow up („one folder up", My Computer)	3,53	3,53
Re-read actual site (Browser)	3,31	3,47
Stop loading (Browser)	2,79	2,98
Enter URL address through the keyboard (Browser)	4,02	4,05

Favorites, Bookmarks	3,55	3,88
New register tab (Browser)	3,95	2,94
New window (Browser)	3,78	3,53
Search/find for text on the screen (Browser, Docs)	3,35	3,88
Save/open image and/or location (Browser)	3,59	3,19
Print	3,41	2,9
Cut	4,11	3,88
Paste	4,26	4,41
Copy	4,49	4,56
Move	4,26	4,38
Delete	4,14	4,56
New folder (My Computer)	4,24	4,38
Download mails/open E-Mail client (Mail)	4,1	4,59
Compose, create new mail (Mail)	4,2	4,67
Reply (Mail)	4,14	4,67
Reply all… (Mail)	2,91	2,98
Forward (Mail)	3,26	4,19
Save mail/drafts (Mail)	3,18	3,47
Send (Mail)	4,22	4,67
Address book (Mail)	3,51	3,72
Attachment (Mail)	3,7	4,06
Open	4,43	4,71
Save	4,28	4,71
Save as…	4,29	4,71
Close	4,27	4,65
Rename	4,14	3,94
Restore from Recycle Bin	3,04	3,44
Empty Recycle Bin	3,74	3,83
New Document	4,29	4,53
Spelling (Docs)	3,47	3,41
Font size (Docs)	3,79	3,41
Format: **B**/*I*/U (Docs)	3,98	3,47
Select, mark, highlight text (Docs)	3,91	4,29
Repeat	3,78	2,94
Undo	3,78	2,94
OTHERS		
Waiting… (hour-glass)		
Start, shut-down, restart computer…		
Resize windows (grow, shrink)		
Frame/border of the screen		
Scrolling		
Menu navigation		
Actual time, system clock		
EMOTICONS		

Table 1: Averaged points given by the subjects.

3. EVALUATION OF AUDITORY EVENTS

The approach has been to create sound samples with a length of about 1,5 sec. The actual duration is between 0,6 and 1,7 sec with an average of 1,11 sec. There are two types of sounds: normal sounds that have to be played back once; and sounds to be played back repeated (in loop). The first represents icons, menu items or short events. Looped signals are active and are supposed to be played back during a longer action or event (e.g. copying, printing).

Sound files were recorded or downloaded from the Internet and were edited by Adobe Audition software in 16 bit, 44100 Hz mono wave format [9, 10]. An actual implementation can use a bitrate compressed format as well such as mp3. Editing includes simple operations of amplifying, cutting, mixing, fade in/out effects. At the final stage, all samples have about the same loudness level (±1 dB).

A collection of wave data of about 300 files was categorized, selected and evaluated. Subjects were asked to identify the sound (what is it) and judge them by „comfortability" (how pleasing it is to listen to it). During evaluation subjects ranked different sound samples (types) and variations to a given application or event. For example, it had to be determined how „open" sounds like? Possible sound samples include slide fastener (opening the zip fly on a trouser), opening a drawer, opening a beer can or pulling the curtains. We presented different versions of each to pick the best representation. Furthermore, subjects were asked to think about the sound of „close" – a representation in connection with „open". Therefore, we tried to present reversed versions of opening sounds (simply played back reversed) or using the squeezing sound of a beer can. The reverse playback method can not be applied every time; some samples could sound completely different reversed. Subjects could make suggestions for new sounds as well.

If there was no definite winner or no idea at all, a spearcon version was used (e.g. for Acrobat).

The sound files listed in Tables 2-5 (right columns) are included in a ZIP file that can be directly downloaded from http://vip.tilb.sze.hu/~wersenyi/Sounds.zip.

3.1. Applications

Table 2 shows the most important applications and programs that have to be represented by an auditory event. These were selected if both blind and sighted users ranked them as "important" or "everyday use" (having an average point of at least 3,00 on the questionnaire), except the My Documents folder, and JAWS because these were only important for blind users.

The sound for internet browsing includes two different versions, both were accepted by the users. Interesting is that the sound sample "search/find" contains a human non-speech part that is very similar for different languages, and this is easy to relate to an "impatient human". Subjects could relate the intonation to a possible sentence of "Where is it?" or 'Wo ist es?" (in German) or even "Hol van már?" (in Hungarian). The same intonation is used in different languages to express the feeling during impatient searching. The same sound will be used in other applications where searching, finding is relevant (Browser, Word, Acrobat etc.).

The table does not contain some other interesting samples, such as a modified sound for the E-mail client, where the applied sound is extended with a frustrated "oh" in case there is

no new mail and a happy "oh" if there is a new mail. Since mail clients do have some kind of sound in case of existing new mails this version is not used.

The events related to the recycle bin also have sound events related to the well-known sound effect of the MS Windows "recycle bin.wav". This is used if users empty the recycle bin. We used the same sample in a modified way to identify the icon, opening the recycle bin or restore a file from it. The application identification uses the "paper noise" and a thin can pedal together. Later, the restoration utilizes the paper noise with human caw. The caw shall give the feeling of a false delete earlier.

Application	Description	Filename
Internet browser (1)	Door opening with keys	Browser1
Internet browser (2)	Knocking and opening a door	Browser2
E-mail client	Bicycle and P.O. box	Mail1
Windows Explorer	Spearcon	S_Explorer
My Computer	Computer start-up beep and fan noise	My Computer
Acrobat	Spearcon	S_Acrobat
Recycle Bin	Pedal of a thin can with the recycle bin sound	Pedal
MS Word	Spearcon	S_Word
CD/DVD burning	Burning flame	Burn
Movie/music player (MS MediaPlayer)	Classic movie projector	Projector
Compressors (ZIP, RAR)	Pressing, extruding machine	Press
Virus/Spam killer	Coughing and "aaahhh"	Cough
MSN Messenger	Spearcon	S_Messenger
Control Panel	Spearcon	S_ControlP
My Documents folder on the desktop	Spearcon	S_MyDocs
Search for files etc.	Seeking and searching with human voice (loop)	Search_(loop)
JAWS/TTS/Screen Reader appl.	Spearcon, speech	

Table 2: Collection of the most important programs and applications (MS Windows based). Sound samples can be found under the given names.

For compressors, we had samples of human struggling while squeezing something, e.g. a beer can, but similar sounds appear later by open, close or delete. Similarly, a ringing telephone was suggested to be a good solution for the MSN/Windows Messenger, but this sound is used by Skype already. Finally, two different samples for "Help" were selected: a whispering human noise and a desperate "help" speech-sample. Because Help was not selected as a very important function, and furthermore, the first sample was only popular in Hungary (Hungarian PC environments use the term "whispering" instead of "help" analog to theatrical prompt-

boxes) and the second contains a real English word, these versions were cancelled.

3.2. Navigation and Orientation

The sounds in Table 3 are the most important during navigation and orientation on the screen, first of all for blind persons. Although, blind users do not use the mouse frequently, sometimes it is helpful to know where the cursor is. The movement of the cursor is a looped sound sample indicating that it is actually moving. The actual position or direction of moving could be determined by increasing/decreasing sounds (such as by scrolling) or using HRTF synthesis and directional filtering through headphones [11, 12]. This is not implemented yet. Using this sound together with a "ding" by reaching the border of the screen allows a very quick access to the system tray, the start menu, or the system clock – these are placed bottom left and right of the screen.

Other sounds	Description	Filename
Moving the mouse (cursor)	Some kind of "ding" (loop)	Mouse_(loop)
Waiting for… (sand-glass turning)	Ticking (loop)	Ticking_(loop)
User intervention, pop-up window	Notification sound	Notify
Border of the screen	Some kind of "ding"	Ding (Border)
Scrolling	Increasing and decreasing freq.	
Menu navigation	Spearcons with modifications	
System clock	Speech	S_SystemClock
Start menu	Spearcon, speech	S_StartMenu

Table 3: Collection of important navigation and orientation tasks (MS Windows based). Sound samples can be found under the given names.

In case of menu navigation spearcons have been already shown to be a powerful possibility [13]. Modifications on spearcons to represent menu structures and levels different speakers (male, female) or different loudness levels etc. can be used. In case of short words, such as Word, Excel, Cut etc. the use of a spearcon is questionable, since these words are short enough without compression in time. Users preferred the original recordings instead of the spearcons in such cases. We did not investigate deeply where the limit is, but it seems that speech samples with only one syllable and with a length shorter than 0,5 s. are maybe too short for a spearcon.

3.3. Functions and Events

Table 4 contains the most important and frequently used sub-functions in several applications. Second column indicates where we can find the given function and we can also see some common visual representations (icons). The sounds related to internet browsing have something to do with "home". Users liked very much the home button represented by a doorbell and a barking dog together – something that would happen if we arrive home. Furthermore, arrows forward, back and up also have something to do with a car: start-up, reverse or engine RPM boost. So all the arrows are linked somehow to a car. Similarly, mailing events have stamping and/or bicycle bell sounds representing a postman's activity.

Bookmarks/favorites in a browser and the address book/contacts in the e-mail client share the same sound of a book, turning pages and a humming human sound. This is another good example for using a non-speech human voice sample interacting with a common sound and thus creating a better understanding and mapping.

The sound for printing can be used in a long version or looped in case of printing (in the background this can be more quiet) or as a short sound event to represent the printing icon or command in a menu. The same is true for copy: a longer version can be used indicating the progress of the copying action (in the background), and a shorter to indicate the icon or a menu item.

The sound for "paste" is one of the most complex samples. It uses the sound of painting with a brush on a wall, a short sound of a moving paint-bucket and the whistling of the painter creating the image of a painter "pasting" something.

In case of "move" there are two versions: the struggling of a man with a wooden box, and a mixed sound of "cut and paste": scissors and painting with a brush.

Based on the comments, the action of saving has something common with locking, securing, so the sound is a locking sound of a door. As an extension, the same sound is used for "save as" with an additional human "hm?" sound indicating that the securing process needs user interaction: a different filename to enter. Opening and closing is very important almost for every application. The sounds have to be somehow related to open and close something and they have to be in pairs. The most popular version was a zip fly of a trouser to open up and close. The same sound that was recorded for opening was used for closing as well: it is simply reversed playback. The increasing and decreasing frequency should deliver the information. The other sample is opening a beer can and squeezing it in case of closing.

3.4. Auditory Emoticons

Table 5 contains the auditory emoticons together with the visual representations. Smileys have the goal to represent emotional content and some of them appear to be similar. As smileys try to represent emotions in an easy but limited (graphical) way, the auditory emoticons also try the same using a brief sound. As in real life, some of them express similar feelings. Summarized, we can say that auditory emoticons:
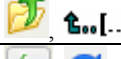- reflect emotional status of the speaker
- are represented always with human sounds, non-verbal and language independent
- can also contain other sounds, noises etc. for a deeper understanding.

Although there is no scientific evidence that some emotions can be represented better by a female voice than by a male voice, we observed that subjects prefer the female version for smiling, winking, mocking, crying and kissing. Table 5 contains both female and male versions.

3.5. Presentation Methods

All the auditory representation presented above can be played back as follows:
- in a direct mapping between a visual icon or button: the sound can be heard as the cursor/mouse is on the icon/button, and the auditory event helps first of all the blind users to orientate (to know where they are on the screen).

Events, Functions	Where?	Visual Representations	Description	Filename
Home button	Internet Browser		Doorbell and dog barking	Homebutton
Arrow back	Internet Browser, My Computer		Reverse a car with signalling	Backarrow
Arrow forward	Internet Browser, My Computer		Starting a car	Forwardarrow
Arrow up	My Computer, Explorer		Car engine RPM increasing	Uparrow
Re-read, Re-load actual page	Internet Browser		Breaking a car and start-up	Reread
Typing, entering URL address	Internet Browser		The sound of typing on a keyboard	Keyboard
Open new / close Browser Window	Internet Browser		Opening and closing sound of a wooden window	Window_open Window_close
Search/find text on this screen/document	Internet Browser, E-mail, Documents		Seeking and searching with human voice (loop)	Search_(loop)
Save link or image	Internet Browser		Spearcon	S_SaveImageAs S_SaveLinkAs
Bookmarks, Favorites	Internet Browser		Turning the pages of a book with human sound	Book
Printing (action in progress)	Everywhere		Sound of a dot-matrix printer	Print
Cut	Documents, My Computer, Browser		Cutting with scissors	Cut
Paste	Documents, My Computer, Browser		Painting with a brush, whistle and can chatter	Paste
Copy	Documents, My Computer, Browser		Sound of a copy machine	Copy_(loop)
Move	Documents, My Computer, Browser		Wooden box pushed with human struggling sound or cutting with a scissor and pasting with a brush	Move1 Move2
Delete	Documents, My Computer, Browser		Flushing the toilet	Delete
New Folder…	My Computer		Spearcon	S_New
New mail, create/compose new message	E-mail		Breathing and stamping	Composemail
Reply to a mail	E-mail		Breath and stamp (once)	Replymail
Forward mail	E-mail		Movement of paper on a desk	Forwardmail
Save mail	E-mail		Sound of save and bicycle bells	SaveMail
Send mail	E-mail		Bicycle bell and bye-bye sounds	Sendmail
Addressbook	E-mail		Turning the pages with human sound	Book
Attachment to a mail	E-mail		Stapler	Attach
Open	Documents, Files		Zip fly up or opening beer can	Zip_up Beer_up
Save	Documents, Files		Locking a door with keys	Save
Save as…	Documents, Files		Locking a door with keys with human "hm?"	Save_as
Close	Documents, Files		Zip fly down or squeezing beer can	Zip_down Beer_down
Rename	Documents, Files		Spearcon	S_Rename

Restore from the recycle bin	Recycle bin		Original "paper sound" of MS Windows and human caw	Recycleback
Empty recycle bin	Recycle bin		Original sound of MS Windows (paper sound)	Recycle
New Document (create)	Documents		Spearcon	S_New
Text formatting tools	Documents		Spearcons	S_Fontsize S_Formatting S_Bold S_Italic S_Underline S_Spelling
Mark/select (text)	Documents, Browser, E-mail		Sound of magic marker pen	Mark

Table 4: Collection of the most important actions and functions (MS Windows based). Sound samples can be found under the given names.

Auditory Emoticon	Visual Representation	Description	Filename (Female)	Filename (Male)
Smile	☺, :-), :),	chuckle	Smile_f	Smile_m
Laughter	:-D	laughing	Laugh_f	Laugh_m
Wink	;-)	Short "sparkling" sound and chuckle	Wink_f	Wink_m
Mock (tongue out)	:-P	Typical sound of tongue out	Tongue_f	Tongue_m
Surprise	:-o	"oh"	Surprise_f	Surprise_m
Anger	,	"grrrrrrrr, uuuhhh"	Anger_f	Anger_m
Perplexed, distracted	:-S,	"hm, aaahhh"	Puzzled_f	Puzzled_m
Shame, "red face"		"iyuu, eh-eh"	Redface_f	Redface_m
Sadness, sorry	☹, :-(, :(,	A sad "oh"	Sad_f	Sad_m
Crying, whimper		Crying	Cry_f	Cry_m
Kiss	:-*,	Sound of kiss on the cheek	Kiss_f	Kiss_m
Disappointment	:-I,	"oh-hm"	Dis_f	Dis_m

Table 5: Collection of the most important emoticons. Sound samples for female and male versions can be found under the given names.

- during an action in progress, e.g. during copying, deleting, printing etc. in loop.
- after an action is finished and completed as a confirmation sound.

The sounds have to be investigated further to find which presentation method is the best for a given action and sound. It is possible that the same sound can be used for both: e.g. first, the sound is played back once as the cursor is on the button "back arrow", and after clicking, the same sound can be played back as a confirmation that the previous page is displayed.

3.6. Spearcons

Two male German native speakers recorded the speech samples and a MATLAB routine was used for the creating the spearcons. One set is called accent-free, while the other speaker had a typical Eastern German Saxonian accent. All spearcons are original recordings in an anechoic chamber using Adobe Audition software and Sennheiser microphones. Spectral evaluation and quality requirements of recordings for German and Hungarian spearcons are discussed in [6]. The Hungarian database was recorded by a native male speaker of 33 years of age. Finally, an English database was recorded by the same Hungarian speaker but this was not used for evaluation yet. The databases contain 27 words (spearcons) respectively.

We observed that longer words (having more vowels) are harder to understand after creating the spearcons. Although, it is not required to understand the spearcon, subjects preferred those they have actually understood. Independent of the fact whether we use a spearcon or not, all were tested and judged by the subjects. All 27 spearcons were played back in a random order (both accent-free versions and Saxonian). A spearcon could be identified and classified as follows:

- the subject has understood it the first time,
- the subject could not understand it, and he had a second try,
- if the subject failed twice, the spearcon was revealed (the original recording was shown) and a final try was made.

The evaluation showed that only 12% of the spearcons were recognized at first. It was interesting that there was no clear evidence and benefit for using accent-free spearcons: recognition of the spearcon sometimes was better for the

Saxonian version. Blind persons tend to be better in this task than sighted persons.

4. FUTURE WORKS

Our current investigation was about to find the best fitting auditory representations for each application and function that had been selected earlier. Future works includes implementation into various software environments such as JAWS (Job Access With Speech) or other Screen Readers that also offer non-speech solutions (Fig. 1-2). The pre-defined samples can be replaced and/or extended with these. Furthermore, a MS Windows patch or plug-in is planned to be created. This executable file can be downloaded, extracted and installed. It will include a simple graphical user interface with checkboxes and simple environmental settings (e.g. auto start on startup, default values etc.) and all of the sound samples, probably in mp3 format.

Figure 1. *Training of blind users to access personal computers [14].*

Figure 2. *The Braille input/output module [14].*

5. CONCLUSIONS

Fifty blind and hundred users with normal vision participated in a survey in order to find the most important and frequently used applications, and furthermore, to create and evaluate different auditory representations for them. These auditory events included auditory icons, earcons and spearcons of German and Hungarian language. The German spearcon database contains original recordings of a native speaker and samples with Saxonian accent. Furthermore, a new class of auditory events was introduced: the auditory emoticons. These represent icons or events with emotional content, using non-speech human voices and other sounds (laughter, crying etc). The previously selected applications, programs, function, icons etc. were mapped and sound samples were evaluated based on subjective parameters. Without a detailed presentation of evaluation data and numbers, only the "winning" sound samples were collected and presented here. Both target groups liked and welcomed the idea and representation method to extended and/or replace the visual content of a computer screen.

6. REFERENCES

[1] D. Burger, C. Mazurier, S. Cesarano, and J. Sagot, "The Design of Interactive Auditory Learning Tools," *Non-visual Human-Computer Interaction*, vol. 228, pp. 97–114, 1993.

[2] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and Icons: their structure and common design principles," *Human-Computer Interaction*, vol. 4, no. 1, pp. 11-44, 1989.

[3] M. L. M. Vargas and S. Anderson, "Combining speech and earcons to assist menu navigation," in *Proc. of the 9th Int. Conf. on Auditory Display (ICAD 03),* Boston, MA, USA, pp. 38-41, July 6-9, 2003.

[4] B. N. Walker, A. Nance, and J. Lindsay, "Spearcons: Speech-based earcons improve navigation performance in auditory menus," in *Proc. of the 12th Int. Conf. on Auditory Display (ICAD 06),* pp. 63-68, London, UK, 2006.

[5] D. K. Palladino, and B. N. Walker, "Learning rates for auditory menus enhanced with spearcons versus earcons," in *Proc. of the 13th Int. Conf. on Auditory Display (ICAD 07),* 6 pages.

[6] Gy. Wersényi, "Evaluation of user habits for creating auditory representations of different software applications for blind persons," in *Proc. of the 14th Int. Conf. on Auditory Display (ICAD 08),* Paris, France, June 23-28, 2008, 5 pages.

[7] http://www.w3.org/

[8] http://www.independentliving.com/prodinfo.asp?number=CSH1W

[9] http://www.freesound.org

[10] http://www.soundsnap.com

[11] Gy. Wersényi, "Localization in a HRTF-based Virtual Audio Synthesis using additional High-pass and Low-pass filtering of Sound Sources," *Journal of the Acoust. Science and Tech. of Japan,* vol. 28, no. 4, pp. 244-250, Jul. 2007.

[12] Gy. Wersényi, "On the improvement of virtual localization in vertical directions using HRTF synthesis and additional filtering," in *Proc. of the 19th Int. Congress on Acoustics (ICA 08)*, Madrid, Spain, Sept 2-7, 2007, 5 pages.

[13] T. Dingler, J. Lindsay, and B. N. Walker, "Learnabiltiy of Sound Cues for Environmental Features: Auditory Icons, Earcons, Spearcons, and Speech," in *Proc. of the 14th Int. Conf. on Auditory Display (ICAD 08),* Paris, France, June 23-28, 2008, 6 pages.

[14] http://index.hu/tech/net/latnet081030/