# 8

# EARCONS AND ICONS: AN EXPERIMENTAL STUDY

*Eric Leimann and Hans-Henning Schulze*

Fachbereich Psychologie, Gutenbergstrasse 18, 35032 Marburg/Lahn, Germany
E-Mail: Schulze@mailer.hrz.uni-marburg.de

**KEY WORDS:** graphical user interfaces, blind users, auditory symbols, comparison of earcons and icons, experimental usability study

**ABSTRACT:** In order to provide access to graphical user interfaces for blind computer users, other ways than iconic representation have to be found. In the present study 22 auditory symbols were developed, referred to as earcons, that stand for some frequently used computer operations. The usefulness of these symbols was tested in three experiments, "identification", "assortment" and "assortment after a learning trial" and then compared to icons representing the same operations. Although further research is necessary, results show that earcons can truly be an alternative representation mode.

## 1. INTRODUCTION

While everybody is happy about the advantages of graphical user interfaces, blind people are not. The reasons for this are obvious. In the old text-based systems the characters shown on the screen are stored as ASCII-signs in a text buffer, which can be transformed directly into speech or Braille-output. Graphical user interfaces (GUIs) make use of the powerful visual information processing capability of sighted people. Features are found such as windows, icons and menus. With the help of pointing devices we locate and manipulate those objects in space. In order to provide access to GUI systems for blind users, ways have to be found to translate the graphic objects into tactile or acoustical perceptable forms. Over the last years there have been several ideas and conceptions of how to achieve this (Blattner et al., 1989; Bly, 1985; Edwards, 1989; Gaver, 1989). Some products that mainly focus on extracting textual information from GUIs are already on the market. Ongoing research tries to provide access to the full range of information including texts, graphics, metaphors and the spatial distribution of all screen elements (Crispien and Petrie, 1993; Bölke and Gorny, 1994). The most difficult problem within this field may be located on the semantical level. Information is provided that is presented in a graphical manner. There is a metaphorical relation between most screen objects and the data to be presented. The question is whether it is possible to transform those metaphors so that visually impaired users can work with GUIs as effectively as their sighted colleagues. Textual description of graphics is long, complicated and imprecise. Existing tactile displays only have a small resolution and are very expensive.

Some researchers have proposed acoustical metaphors to illustrate the screen objects. Video-games for example give a good idea of the potential of non-speech audio cues at the interface. It was shown that expert players' scores are much lower with the audio turned off than when it is turned on (Buxton, 1989). William Gaver (1989) has developed an application he calls the SonicFinder, which is an interface for the Apple Macintosh. He provides auditory icons in a work station environment using everyday sounds to offer feedback on common operations such as copying files or dragging windows. Meera Blattner (1989) and her colleagues have analyzed the common design principles of icons and their aural counterparts, which they describe as earcons. They use short musical motives and rhythmic sequences of pitches for which timbre, register and dynamics may vary. With the help of these parameters, a hierachical structure is created that matches the way data is presented on the GUIs.

K. Nordby et al. (eds.), *Human-Computer Interaction*

No matter what kind of auditory representation the researchers propose, until now there has not been sufficient evalution of acoustical symbols. It is claimed that Icons represent the objects and actions of the computer world in an intuitive way. The present study looks at whether earcons can achieve the same. In the experiments described below, we compare auditory with visual symbols to see which representation works better for an exemplary selection of interface operations.

## 2. SELECTION OF OPERATIONS AND SYMBOLS

At first we made up a list of 12 typical interface operations that are carried out on a routine basis by users of GUIs (see table 1). Two more operations were selected especially for auditory interaction. Then a questionnaire was then developed, where the chosen operations were presented to subjects. They were asked to think of possible auditory representations for each of the given items. At the beginning of the questionnaire we introduced the participants to the concept of acoustic symbols at the interface. Although we use the name earcons for those symbols, it does not mean that we intend to work with musical motives as they were introduced by Blattner et al. (1989). At that point of time we did not have a preference for a special sound concept. The ideas of 67 subjects were categorized and counted. Most of the ideas were examples for what Gaver (1989) describes as everyday sounds. He points out that our normal mode of hearing is to listen to sounds to identify the events that cause them. Only when identification of source events is not possible, we tend to describe sensory qualities of sounds. There were two criteria for selecting an idea for sound realization or not: First, a large number of subjects proposing this idea for a particular operation and second, the idea had to be specific for this operation.

As a result of this selection, we recorded 22 sound events, most of them being natural sounds of our daily environment (see table 1). With the help of an Akai S3000 sampler, some sounds were "time stretched", which means that we shortened the event without changing the pitch. All earcons have a length between 500 and 1500 ms (mean duration: 1070 ms). No other manipulations of the original recordings have been carried out. As mentioned above, we also wanted to compare the earcons with icons widely used in graphical user interfaces. We chose visual symbols from the GUI *Microsoft Windows (MS)* and current versions of three text processing programs: *Word for Windows 2.0* (WfW), *Word Perfect for Windows 5.1* (WP) and *Ami Pro 3.0* (Ami).

Table 1: Operations and symbols

| **operations** | **earcons** | **icons** |
|---|---|---|
| create | cry of a newborn baby<br>insert paper into a typewriter<br>start a car | WfW<br>Ami |
| save | lock (a door) | WfW<br>Ami<br>WP |
| load | load a gun<br>gunshot and reload | WfW<br>Ami<br>WP |
| print | typewriter<br>wire printer | WfW<br>Ami<br>WP |
| delete | toilet flush<br>crumble a paper<br>douse a fire<br>tear a paper | Ami |
| select | mark with a texter | |
| move (paste) | drag a chair across the floor | WfW |
| copy | click of a camera<br>noise of a photocopier | WfW<br>WP |
| search | snuffling dog | Ami<br>WP |
| exit / quit | shut a door | Ami |
| escape | running footsteps | Ami<br>WP |
| open window (pop up) | opening a window (traffic) | |
| close window | close a window (noise stops) | MS |
| help | sound of exclamation | Ami |

### 3. EXPERIMENT 1: IDENTIFICATION OF EARCONS

The purpose of this experiment was to investigate how well the presented sounds could be identified. The fact that a sound is identified by most subjects as the same specific event should be an important requirement for a powerful symbolism. Forty subjects were recruited from the student population of the University of Marburg. All subjects had either normal or corrected-to-normal vision. Due to the fact that no evidence exists showing that seeing and blind individuals identify or imagine auditory symbols differently, visually impaired people were not included in the sample. Because the participants were non-computer users, they had not seen the icons before. Although the earcon concept was explained to them, they knew nothing about the operation-symbol relation. Subjects were instructed to write down a short description of the sound they heard. The presentation mode was the same for all three experiments. Sound samples were played by a digital audio tape recorder (DAT) and presented through headphones. First, subjects heard all 22 earcons in a rapid and random succession. There was a sound event every two seconds with every earcon being presented twice in succession. After a short break of ten seconds, a voice announced "sound one". Two seconds later "sound 1" was presented and repeated after four more seconds. Subjects then had 14 seconds time to carry out their task before "sound 2" was announced. One complete presentation took a total of ten minutes. The percentage of correct identifications is given in the column "Ident." of table 2. Because a correct identification of an icon can only be defined for icons representing specific objects table 2 does not include a column "Ident." for icons.

Table 2: Percentage of correct responses by symbol identification, assortment and assortment after a learning trial

| Earcons | Ident. | Assort. | Learn. | Icons | Assort. | Learn. |
|---|---|---|---|---|---|---|
| cry of a newborn baby | 100% | 0% | 72% | WfW-create | 46% | 62% |
| insert paper into typewriter | 80% | 6% | 30% | Ami-create | 30% | 65% |
| start a car | 35% | 8% | 7% | WfW-save | 42% | 47% |
| lock a door | 63% | 9% | 67% | Ami-save | 32% | 64% |
| load a gun | 5% | 20% | 14% | WP-save | 20% | 24% |
| gunshot and reload | 33% | 3% | 4% | WfW-load | 33% | 52% |
| typewriter | 100% | 42% | 83% | Ami-load | 24% | 48% |
| wire printer | 75% | 74% | 72% | WP-load | 10% | 71% |
| toilet flush | 83% | 79% | 91% | WfW-print | 85% | 95% |
| crumble a paper | 85% | 65% | 67% | Ami-print | 89% | 100% |
| douse a fire | 18% | 10% | 61% | WP-print | 12% | 33% |
| tear a paper | 63% | 6% | 26% | Ami-delete | 83% | 95% |
| mark with a texter | 0% | 32% | 71% | WfW-move | 16% | 50% |
| drag a chair across the floor | 43% | 50% | 55% | WfW-copy | 81% | 86% |
| click of a camera | 85% | 32% | 56% | WP-copy | 15% | 78% |
| noise of a photocopier | 28% | 49% | 88% | Ami-search | 42% | 67% |
| snuffling dog | 73% | 47% | 88% | WP-search | 24% | 84% |
| shut a door | 80% | 41% | 60% | Ami-exit | 53% | 76% |
| running footsteps | 98% | 26% | 42% | Ami-escape | 7% | 33% |
| opening a window (traffic) | 30% | 54% | 75% | WP-escape | 11% | 14% |
| close a window (noise stops) | 8% | 55% | 63% | MS-Windows | 39% | 65% |
| sound of exclamation | 100% | 78% | 80% | Ami-help | 77% | 100% |
| **Mean** | **58%** | **36%** | **58%** | | **40%** | **64%** |

## 4. EXPERIMENT 2: ASSORTMENT OF SYMBOLS AND OPERATIONS

The objective of this experiment was to investigate how well the auditory and visual symbols could be associated with the computer operations they represented. The sample was again 40 students, not identical with those of Experiment 1, but from the same population. After an introduction to the icon/earcon concept, subjects received a list with 12 or 14 computer operations (see table 1), including a definition for each item. In the case that a subject did not understand the meaning of a definition, the instructor helped with further explanations. The presentation of the auditory stimuli was as described in Experiment 1. The 22 numbered icons were presented on a sheet of paper. The subjects' task was to assort the number of a given sound or picture with the operation it stands for. Subjects were told that each symbol stands for one operation, but that one operation can have several symbols. Like in Experiment 1, the order of presentation for sound and picture blocks was varied over the subjects with each assortment task lasting ten minutes. For both symbol groups, the percentage of correct assortments is given in table 2.

Table 3: Comparison of the best earcon and icon for each operation (assortment)

| Operation | Best Icon | Best Earcon |
|---|---|---|
| print | Ami (89%) | wire printer (74%) |
| move | WfW (16%) | drag a chair (50%) |
| close window | MS (39%) | close win. (55%) |
| help | Ami (77%) | exclamation (78%) |
| copy | WfW (81%) | photocopier (49%) |
| load | WfW (33%) | load a gun (20%) |
| delete | Ami (83%) | toilet flush (79%) |
| create | WfW (46%) | start a car (8%) |
| exit/quit | Ami (53%) | shut a door (41%) |
| escape | WP (11%) | footsteps (26%) |
| save | WfW (42%) | lock a door (9%) |
| search | Ami (42%) | snuffl. dog (47%) |
| **Mean** | **51.0%** | **44.7%** |

The difference of number of correct judgements for icons and earcons was not statistically significant (t=.054, df= 34). In a more detailed analysis, we compared the best earcon - in terms of number of correct judgments- and best icon for each operation. Table 3 shows the percentages of correct assignments for each operation.

## 5. EXPERIMENT 3: ASSORTMENT AFTER A LEARNING TRIAL

In the final experiment we wanted to examine the effect of a two minute learning period on the result of the assortment task. Twenty-five subjects (same population) participated. The icons were presented on a sheet of paper with the correct operation names written next to the picture. During the next trial, subjects listened to the earcons numbered 1 to 22, with every sound being played twice. Subjects could read the corresponding operation names on a sheet of paper lying in front of them. After a two minute break the assortment task for the first learning trial began. The assortment task was the same as in Experiment 2 but with only eight minutes time. Finally, the second assortment task had to be carried out. Again, the presentation order of icons and earcons was varied over subjects. A global comparison of correct assignments for earcons and icons after learning revealed that performance for the icons was significantly better than for the earcons (t=2.3, df=20, p=.03).

Table 4: Comparison of the best earcon and icon for each operation (assortment after learning trial)

| Operation | Best Icon | Best Earcon |
|---|---|---|
| print | Ami (100%) | typewriter (83%) |
| move | WfW (50%) | drag a chair (55%) |
| close window | MS (65%) | close win. (63%) |
| help | Ami (100%) | exclamation (80%) |
| copy | WfW (86%) | photocopier (88%) |
| load | WP (71%) | insert paper (30%) |
| delete | Ami (95%) | toilet flush (91%) |
| create | Ami (65%) | cry (baby) (72%) |
| exit/quit | Ami (76%) | shut a door (60%) |
| escape | Ami(33%) | footsteps (42%) |
| save | Ami (64%) | lock a door (67%) |
| search | WP (84%) | snuffl. dog (88%) |
| **Mean** | **74.1%** | **68.3%** |

The detailed comparison at the level of the operations is shown in table 4.

## 6. THE RELATION BETWEEN IDEA, IDENTIFICATION, ASSORTMENT AND LEARNING

For each earcon we can define four variables: The proportion of subjects that proposed the specific idea in the selection phase of the study, the proportion of correct identifications, intuitive assignments and assignments after learning. The correlation between these variables is shown in table 5.

Table 5: Relation between idea, identification, assortment and assortment after learning trial (Pearson correlation matrix)

| | Idea | Ident. | Assort. |
|---|---|---|---|
| **Idea** | | | |
| **Ident.** | -0.312 | | |
| **Assort.** | 0.289 | 0.182 | |
| **Learn.** | 0.449 | 0.289 | 0.658 |

## 7. DISCUSSION

When comparing visual and auditory symbols, one must admit that people are much more used to the first. Our daily environment is full of icons. They are on road signs, microwave ovens and directions for building, only to name a few examples. The use of everyday sounds as symbols on the other hand, should have been quite a new thing for our subjects. With this background we found the earcon results of the assortment and assortment after learning trial quite remarkable. In the global comparison of correct assignments, the earcons were with 36% versus 40% for assortment, and 58% versus 64% for assortment after learning trial only slightly less successful than the icons. Of course the reported experiments can only be considered a beginning. The pool of presented earcons was much to small to make global assumptions. The general significant advantage of icons in Experiment 3 could therefore possibly be due to a few bad earcons in the pool (e.g. "start a car" or "gunshot and reload"). On the other hand, the comparison at the level of operations shows that some of the 22 presented earcons work much better as a symbol for an operation than the widely used icons (e.g. Experiment 2: "drag a chair" for -move- or "running footsteps" for -escape-). The low percentage of correct assignments for some symbols shows the importance of a well-considered selection and evaluation of both earcons and icons.

The analysis of relations between idea, identification, assortment and assortment after a learning trial provides us with new insights. Of course there is a high correlation (.66) between assortment and assortment after learning trial. Earcons that are intuitively related to a specific operation should be learned easily. The relation between identification and assortment (.18), as well as between identification and assortment after a learning trial (.29), is positive but not as strong as one might expect. Isn't it necessary to identify a sound correctly to use it as an efficient symbol? Let us look at an example: The sound of "mark with a texter" was not identified at all in Experiment 1 (0%). Surprisingly, this "mysterious" earcon got quite good results in Experiments 2 (32%) and 3 (71%). "Closing a window" (8% - 55% - 63%) is another example for this phenomenon. One possible explanation could be that there are sounds that, though their sources can not be identified, have attributes that can be related to a specific operation. An analysis of subjects' mistakes in the identification task could be helpful to answer this question. The example of "cry of a newborn baby" (100% - 0% - 72%) shows another interesting relation. Here we have an earcon that is identified perfectly but can not be related to the operation that it should symbolize. However, once subjects are told this relation, they can remember it quite well. Finally, we want to take a look at the relation of ideas and experiments. Correlations of .29 (idea - assortment) and .45 (idea - learning trial) indicate that ideas for earcons that are proposed by many subjects might have good results in an intuitive assortment task and seem to be successful within a learning task.

## REFERENCES

Blattner, Meera M., Sumikawa, Denise A. & Greenberg, Robert (1989). Earcons and Icons: Their Structure and Common Design Principles. *Human Computer Interaction, 4, 1989, 11-44.*

Bly, Sara (Ed.) (1985). Communicating with sound. *Proceedings of CHI'85 Conference on Human Factors in Computer Systems, New York: ACM, 115-119.*

Bölke, Ludger, Gorny, Peter (1994). Auditory direct manipulation of acoustical objects by blind computer users. *Report, University of Oldenburg, Germany, - Informatics Department*

Buxton, William (1989). Using Nonspeech Audio: An Introduction. *CHI'89, The Use of Non-Speech Audio at the Interface, Tutorial Paper, 1.1-1.6.*

Crispien, Kai, Petrie, Helen (1993). Providing Access to GUIs for Blind People Using a Multimedia System Based on Spatial Audio Presentation. *Audio Engineering Society 95th Convention, New York, Oct 1993.*

Edwards, Alistair D.N. (1989). Soundtrack: An Auditory Interface for Blind Users. *Human Computer Interaction, 4, 1989, 45-66.*

Gaver, William W. (1989). The SonicFinder: An Interface That Uses Auditory Icons. *Human Computer Interaction, 4, 1989, 67-94.*