In B. Blumenthal, J. Gornostaev & C. Unger (Eds.), Human Computer Interaction (Lecture Notes in Computer Science, Vol. 876, pp. 35-44). Berlin: Springer.

Positive Effects of Sound Feedback during the Operation of a Plant Simulator

Matthias Rauterberg and Erich Styger

Usability Laboratory, Work and Organisational Psychology Unit,
Swiss Federal Institute of Technology (ETH)
Nelkenstrasse 11, CH-8092 ZURICH
Tel: +41-1-632-7082, email: rauterberg@rzvax.ethz.ch

Abstract

An experiment was carried out to estimate the effect of sound feedback on the work of a plant operator. Eight students of computer science operated a process simulation program of an assembly line with computer numeric controlled (CNC) robots. Relevant information of disturbances and machine breakdowns was given only in a visual (test condition 1), and in visual and audible form (test condition 2). The results indicate, that the additional sound feedback improves significantly the operator performance and increases positively some mood aspects.

KEYWORDS: Audible feedback, non speech sound generation, human - computer interaction, visual strain.

1 Introduction

Using non speech sounds to provide system information is appealing for several reasons. First, by adding sound to the interface the bandwidth of communication can be significantly increased. Second, the information conveyed by sounds is complementary to that available visually, and thus sound can provide a mean for displaying information that is difficult to visualise, especially with limited screen real estate. Sound feedback can help to improve the usability of interfaces in the following ways:

User Interface Design – Most of all user interfaces stresses the visual perception. Sound feedback can help to reduce eye strain.

Multimedia – New possibility for the interactive representation of complex sound generating events and processes.

User interfaces for people with impaired vision – Simulations with the utilisation of audio data will in future also have their application in the training of people with impaired senses, in particular of people with damaged vision.

The hearing of sounds in everyday life is based on the perception of events and not on the perception of sounds as such [Rauterberg-94]. For this reason, everyday sounds are often described by the events they are based on. Sound is a familiar and natural medium for conveying information that we use in our daily lives. The following examples help to illustrate the important kinds of information that sound can communicate [Mountford-90]:

Information about physical events – We can hear whether a dropped glass has bounced or shattered.

Information about invisible structures – Tapping on a wall is useful in finding where to hang a heavy picture.

Information about dynamic change – As we fill a glass we can hear when the liquid has reached the top.

Information about abnormal structures – A malfunctioning engine sounds different from a healthy one.

Information about events in space – Footsteps warn us of the approach of another person.

The textual representation of information is of most use when the user is familiar with the domain area and can demonstrate much experience and knowledge in that domain area [Marmolin-92]. In comparison, more concrete (visual and auditory) representations of information that the user can query are of most use when the domain area is new and unknown.

The parallel use of different media and the resulting parallel distribution of information, for example by simultaneously showing a predecessor through a concrete representation and its explanation through audio distribution, leads to a denser sharing of information. In this case, the user can dedicate his attention solely to the visual information, which has parallel audio support. This reduces the need to change the textual or other visual delivery and prevents the overflow of visual information [Edwards-88].

Sounds and music can be utilised to improve the user's understanding of visual predecessors or can stand alone as independent sources of information. Gaver et al [Gaver-91] used sounds as diagnostic support applied with the direction of a process simulation in a collaborative environment. But, he did not prove the hypothesis that an interface with sound is superior to an interface without sound feedback. Gaver et al [Gaver-91] describe only some global impressions of different user reactions to sound feedback.

Our main interest was to test the hypothesis of [Buxton-89] and [Gaver-91], that people in the real world monitor multiple background activities simultaneously through sound. So, we use auditory cues to help users to monitor the status of ongoing processes. Diagnosing and treating problems with the plant were aided by alert sounds (see also [Gaver-91]). In difference to [Gaver-91] we used individual sessions, and not a collaborative environment. We carried out an experiment, that allows us to test our hypothesis with the methodology of applied statistics.

2 Method

2.1 Subjects

Eight male students of computer science at the ETH took part in the experiment (mean age of 24 ± 1 years).

2.2 Simulator

The simulation is based on a flexible manufacturing system, that produces cases made of aluminium (see 'work pieces' in Figure 1). The whole system consists of eight computer-numeric-controlled (CNC) manufacturing centres and eight loading robots for these centres. In the input directing station all work pieces are automatically directed on the assembly line. The assembly line transports each work piece through different stations to the CNC manufacturing centres and back to the output directing station. The whole plant was deliberately designed to be too large to fit on the computer screen, so users could only see about half the CNC machines at any time (see 'actual screen clipping' in Figure 1).

We designed our simulator so that each of the machines made sounds to indicate its status over time. Each sound was designed to reflect the semantic of the actual event. For instance, a splashing sound indicated that cooling liquid was being spilled. Because of the complexity of our system, as many as 38 sounds made be placed at once. We attempted to design the sounds so that none would be masked (rendered inaudible) by other sounds. [Gaver-91] describe two strategies to be useful in avoiding masking. First, sounds were spread fairly evenly in frequency, so that some were high-pitched and others lower. Second, we avoided playing sounds continuously and instead played repetitive streams of sounds, thus maximising the chance for other sounds to be heard in the gaps between repetitions. CNC 0 and CNC 4 are characterised by a high-pitched sound. CNC 3 and CNC 7 are low-pitched.

A work piece could have one of the following status: (1) loading on the assembly line at the input directing station, (2) transportation on the assembly line, (3) fixation on the carrier at the reset station, (4) final fixation and twist on the carrier, (5) fixation on a pallet with three other work pieces at the robot, (6) processing one of two sides in the CNC station, (7) change from one side to the other at the reset station, (8) to be provided with a serial number at the labelling station, (9) loading off the assembly line at the output directing station. Steps (3) to (7) are carried out twice, once for each side of the work piece.

Table 1: Sound types, duration, and size.

machine	sound	duration	size
CNC 0-7	normal	1.20 s	51 KB
CNC 0-7	no cooling	1.08 s	46 KB
CNC 0-7	jammed pipe	1.38 s	59 KB
robot 0-7	normal	0.39 s	16 KB
robot 0-7	lost piece	1.04 s	44 KB
robot 0-7	tear off pipe	1.04 s	44 KB
input station	normal	0.41 s	17 KB
output station	normal	0.78 s	33 KB
reset station	normal	1.40 s	60 KB
twist station	normal	0.40 s	17 KB
labelling station	normal	0.49 s	21 KB
control station	global alarm	0.24 s	10 KB

Normal running of a machine was coupled with a characteristic sound pattern. Each machine breakdown generated instead of the normal sound a specific alert sound (see Table 1). If a robot or a CNC centre breaks down, then this centre can not process the pallet of four work pieces further on. The first consequence of a breakdown is a jam on the assembly line. The second consequence is the productivity of the plant decrease.

2.3 Task

Subjects were instructed to operate a plant simulator and to take care for a high productivity rate. The task was to trouble-shoot the whole manufacturing system. First, each subject had to detect that a breakdown happened. Then he has to find the interrupted machine (robot or CNC machine). The actual breakdown event shows the operator how to repair the machine. The operator can get this information visually in a modal dialogue box with the status report at the control station or in an auditory form through sound feedback.

A CNC machine could have two breakdown events ('jammed outlet pipe of cooling agent', 'empty cooling agent'). A robot could breakdown with two different events ('lost work piece', 'tear off a pressure pipe').

Table 2: Machine, breakdown type, and repair code.

machine	breakdown	code
CNC 0-7	no cooling	3713
CNC 0-7	jammed pipe	8319
robot 0-7	lost piece	1731
robot 0-7	tear off pipe	1733
control station	status report	8700

Each interrupted machine could be repaired by entering an appropriate repair code (a four-digit number, see Table 2) in a repair dialogue box at the machine. The operator sees only a part of the whole plant (see 'actual screen clipping' in Figure 1). He moves the actual screen up and down by clicking with the mouse in the scrollbar area to 'go to' the interrupted machine. A mouse click on the machine symbol pops up the repair dialogue box. Entering the correct repair code transfers the interrupted machine in the normal state. If an incorrect repair code is entered, then no internal state change happens and the user could hear only a short beep.

Users' view of the plant behaviour was that robots and CNC centres breakdown accidentally. Our simulation program was programmed so, that all breakdowns appeared in the same sequence. This approach guarantees that the trials between users are maximally comparable.

2.4 Procedure

We run the experiment with a two-factorial test design. Factor A was 'with' or 'without' sound feedback. Test condition 1 was only visual feedback with a warning flasher and a modal dialogue box with status information of each manufacturing system at the operator control station. Test condition 2 was visual and audible feedback of each machine breakdown.

Factor B was a repeated measurement design. Four subjects started the experiment with sound feedback (test condition 1) and repeated the same task without sound feedback (test condition 2). The other four subjects started without sound feedback (test condition 2), and repeated the task with sound feedback (test condition 1).

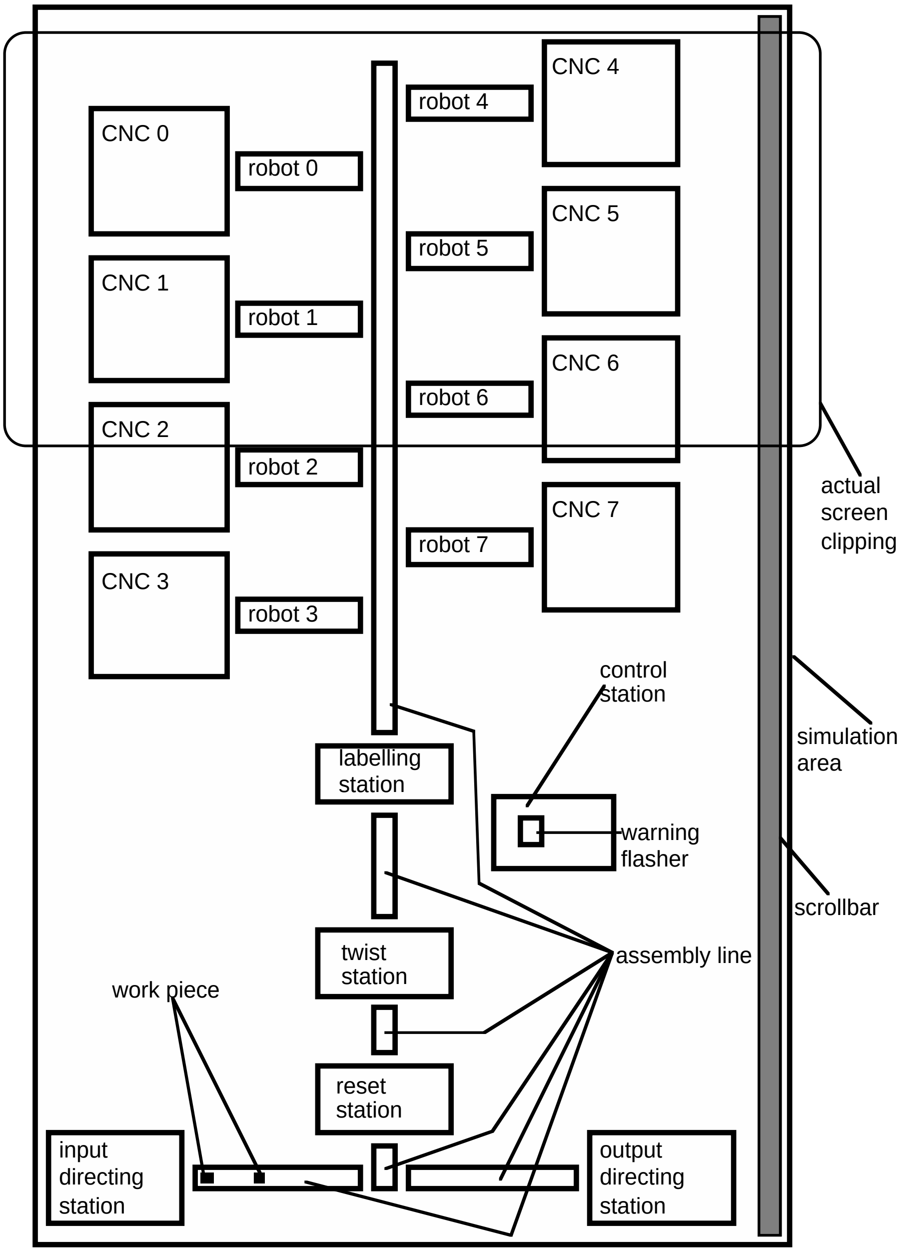

Figure 1: The schematic view of the plant simulator shown to the users. The rectangle shows the actual screen output each user sees at a given time. The whole system consists of eight computer-numeric-controlled (CNC 0..7) manufacturing centres and eight loading robots (robot0..7). In the input directing station all work pieces are automatically directed on the assembly line. The whole plant was deliberately designed to be too large to fit on the computer screen, so users have to scroll.

Each subject filled out a questionnaire to estimate the individual experiences with computers (about 10 minutes). The subjects were introduced in operating the simulation tool through 'learning by using' (about 15 minutes). The simulation ran for the trouble-shooting task exactly 20 minutes. Before and after each trouble-shooting task the user has to answer a mood questionnaire (eight scales with overall 36 items as monopolar rating scales [Apenburg-86]). After each trouble-shooting task we measured the subjective satisfaction with a semantic differential (11 bipolar items). Each individual session took about 90 minutes.

2.5 Material

We ran the experiment on an IBM compatible PC (Olivetti® i386, 25 MHz, 6 MByte main storage, 17" VGA colour screen) with an extra sound card (Logitech® 16 Bit, 44 kHz, stereo). A special simulation program was developed in Turbo Pascal® 1.0 to present the signals on the screen. Users heard the sound out of two small active speakers (maximal 3 watt). All machines at the left side (see Figure 1) could be heard out of the left speaker. The right speaker gave out the sound of all machines at the right side.

2.6 Measures

Our first dependent variable is a point scale that measures the productivity of the plant. Each work piece, that entered the assembly line at the input direction station, counts one point. One point is counted for each side, that was processed at a CNC machine. Each work piece, that left the assembly line at the output direction station, counts an extra point. Each work piece on the assembly line counts one to four points. The productivity score after 20 minute's simulation time is the sum over all work pieces that entered the assembly line.

The second dependent variable is the number of requested status reports at the control station.

The third and fourth dependent variables are number of correct and number of incorrect repairs.

The eight scales of the mood questionnaire and the 11 items of the semantic differential are dependent variables to measure users' satisfaction.

3 Results

First, we present the results of the four dependent variables that measure users' trouble-shooting activities. We find a significant difference between the two test-conditions for the productivity score (see Table 3).

Without sound feedback users moved to the control station and requested the status report significantly more than in the test condition with sound feedback (see Table 3). We could observe, that most of the users in test condition with sound go first to the control station to look for all breakdowns, and go after that through the whole plant to repair machine by machine. During this walk-through they could remember all not repaired machines listening to the sound pattern.

On one side, we can observe a significant improvement through sound feedback, on the other side we can find, that users perceive the simulation with sound more non-transparent and feel more confused than without sound (see Table 4).

Table 3: Results of the four dependent variables that measure users' trouble-shooting activities.

Variable	With sound	Without sound	P sign
productivity score	70 ± 5.6	65 ± 5.3	.052
# of status reports	17 ± 5.8	23 ± 4.0	.043
# of correct repairs	36 ± 2.5	36 ± 2.3	.550
# of incorrect repairs	16 ± 11.0	9 ± 7.1	.113

Table 4: Results of the eleven items of the semantic differential (bipolar rating scale: -2, -1, 0, +1, +2).

Variable (-)................(+)	With sound	Without sound	P sign
time consuming.......time saving	-1.1 ± 0.7	-1.0 ± 0.9	.791
rigid.........flexible	-0.9 ± 1.3	-0.8 ± 0.8	.735
circumstantial..simple	+0.5 ± 2.3	+0.4 ± 3.1	.889
intransp...transparent	+0.4 ± 1.1	+1.4 ± 0.6	.064
confuse....unequivocal	+0.1 ± 2.7	+1.1 ± 1.0	.179
unclear..........clear	0.0 ± 2.6	-0.4 ± 1.4	.596
complicated..uncomplic	0.0 ± 1.1	-0.3 ± 1.9	.712
prescribed........free	-0.5 ± 0.9	-0.4 ± 1.1	.816
unforesee..foreseeable	0.0 ± 2.3	+0.1 ± 1.8	.871
unsuscept..susceptible	-0.8 ± 1.1	-0.9 ± 1.0	.781
angry.........pleasing	-0.4 ± 1.7	-0.1 ± 1.3	.709

Table 5: Results of the differences (after - before) of the eight scales of the mood questionnaire (monopolar rating scale).

Variable	With sound	Without sound	P sign
readiness of endeavour	+2.4 ± 4.1	-0.5 ± 4.1	.199
restfulness	+1.3 ± 2.7	+0.4 ± 3.3	.589
readiness for contacts	+0.9 ± 2.5	-0.8 ± 2.2	.219
drowsiness	-1.1 ± 2.4	-1.5 ± 3.2	.801
self-assurance	+1.8 ± 2.0	-0.6 ± 1.7	.022
social acceptance	+0.1 ± 1.0	-1.1 ± 1.0	.031
to feel excited	0.0 ± 6.1	-1.0 ± 5.9	.738
mood-laden	+1.3 ± 2.2	-0.3 ± 1.0	.128

Users felt significantly more self-assure and more social accepted after working with sound feedback than without sound (see Table 5). Their readiness for endeavour, restfulness, and mood increased in the test condition with sound.

The results of this experiment showed, that the performance of operating a plant simulator could be significantly improved, when feedback of machine break downs and other disturbances was given in an auditory form, too. We can also observe a significant increase of different aspects of users' mood. Overall, we can say that users feel better and less stressed with sound feedback, than without sound.

We found that sound feedback was effective in the following way. Sound feedback helped users keep track of the ongoing processes. Sounds allowed users to track the activity, rate, and functioning of normally running machines. Without sound feedback, users overlooked machines that were broken down. With sound feedback these problems were indicated either by the machine's sound ceasing or by the various alert sounds. Continuous sound feedback allowed users to hear the plant as an integrated complex process. The sounds merged to produce an auditory pattern, much as the many sounds of everyday machines.

4 Discussion

The sense of hearing is an all-round sense. This aspect is an important difference to visual perception, that is a directional sense. An auditory interface can be much larger than the visual interface (screen). Visually hidden aspects of parallel processes in the background can be made perceptible with auditory feedback [Cohen-93]. The results of our experiment support this design approach. Sound feedback of concurrent processes, that are important for task solving, improves the usability of interfaces.

Audition is a spatial sense; we can be aware simultaneously of many sounds coming from different locations. But spatial patterns in audition are much more limited than those of vision. It is primarily a time sense, for its main patterns are those of succession, change, and rhythm. Auditory feedback typically arrives sequentially in time, whereas visual pattern my be presented either sequentially or simultaneously. Of course many perceptual experiences depend on the operation of several senses at once; then the prominence of sense over another becomes a matter for study [Hartman-61].

Sound feedback has poor 'referability', meaning that they usually cannot be kept continuously before the user, although they can be repeated periodically. Visual patterns offer good referability, because the information usually can be 'stored' in the display. The most important advantage of sound feedback is its 'attention-demanding'; it 'breaks in' on the attention of the user. Visually stimuli, however, do not necessarily have this captive audience. The user has to be looking toward the display in order to perceive the stimulus [Rauterberg-93]. Hearing is somewhat more resistant to fatigue than vision [McCormick-57, p. 427].

How many different concurrent sounds can be discriminated? Users reacted up to 38 different sounds in our simulation study. Momtahan et al [Momtahan-93] could show that staff in operating rooms was able to identify only a mean of between 10 and 15 of the 26 alarms. Nurses were able to identify only a mean between 9 and 14 of the 23 alarms found in their intensive care unit. Momtahan et al explain their results with the poor design of auditory warning signals. Standardisation of sound feedback can minimise this perceptual problem.

Cohen [Cohen-93] found that it is a difficult task to design sound pattern "which tell the right story and are also pleasant and emotionally neutral." Good sound feedback needs sound patterns that are interpretable without visual redundancy (e.g., door creaks open, door slams [Cohen-93]). We have to look for everyday sounds that 'stand for themselves'. Given these sounds we have to map them in a metaphorical sense to new events introduced by technology (e.g., door creaks open => login, door slams => logout [Cohen-93]). For simulation tools, that deal with real world events, we can easily use the corresponding real world sounds. The results of our study support this 'real sound' approach.

To avoid disturbances at the workplace we need empirical studies with earphones -- or other possibilities -- to restrict the auditory feedback space to the user's location. Everyday sounds are mostly soft and slight, so, maybe soft sound feedback is an appropriate approach to avoid unintentional effects in a collaborative environment.

5 References

[Apenburg-86| Apenburg, E. (1986) Befindlichkeitsbeschreibung als Methode der Beanspruchungsmessung. *Zeitschrift für Arbeits- und Organisationspsychologie* 30(N.F. 4):3-14.

[Buxton-89] Buxton, W. (1989) Introduction to this special issue on non speech audio. *Human-Computer Interaction* 4(1):1-9.

[Cohen-93] Cohen, J. (1993) "Kirk Here:" Using Genre Sounds To Monitor Background Activity. in S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel and T. White (eds.) INTERCHI'93 Adjunct Proceedings. (pp. 63-64), New York: ACM.

[Edwards-88] Edwards, A. (1988) The design of auditory interfaces for visually disabled users. In E. Soloway, D. Frye and S. Sheppard (eds.) Conference Proceedings "Human Factors in Computing Systems" CHI' 88. (pp. 83-88), New York: ACM.

[Gaver-89] Gaver, W. (1989) The Sonic Finder: an interface that uses auditory icons. *Human Computer Interaction* 4(1):67-94.

[Gaver-91] Gaver, W., Smith, R. & O'Shea, T. (1991) Effective sounds in complex systems: the ARKola simulation. In S. Robertson, G. Olson and J. Olson (eds.) Conference Proceedings "Reaching through Technology" CHI'91. (pp. 85-90), Reading MA: Addison-Wesley.

[Gaver-93] Gaver, W. (1993) Synthesizing auditory icons. S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel & T. White (eds.) Conference Proceedings on "Human Factors in Computing Systems" INTERCHI'93. (pp. 228-235), Reading, MA: Addison-Wesley.

[Hartman-61] Hartman F. (1961) Single and multiple channel communication: a review of research and a proposed model. *Audio-Visual Communication Review* 9(6):235-262.

[Marmolin-92] Marmolin, H. (1992) Multimedia from the perspective of psychology. In L. Kjelldahl (ed.) Multimedia: Systems, interaction, and applications. (pp. 39-52), Berlin, Heidelberg: Springer.

[McCormick-57] McCormick, E. (1957) Human Engineering. New York: McGraw-Hill.

[Momtahan-93] Momtahan, K., Hetu, R. & Tansley, B. (1993) Audibility and identification of auditory alarms in the operating room and intensive care unit. *Ergonomics* 36(10):1159-1176.

[Mountford-90] Mountford, S. & Gaver, W. (1990) Talking and Listening to Computers. In B. Laurel and S. Mountford (eds.) The Art of Human-Computer Interface Design. (pp. 319-334), Reading, MA: Addison-Wesley.

[Rauterberg-93] Rauterberg, M. & Cachin, C. (1993) Locating the primary attention focus of the user. in T. Grechenig & M. Tscheligi (eds.) Human Computer Interaction. (Lecture Notes in Computer Science, vol. 733, pp. 129-140), Berlin: Springer.

[Rauterberg-94] Rauterberg, M., Motavalli, M., Darvishi, A. & Schauer, H. (1994) Automatic sound generation for spherical objects hitting straight beams based on physical models. In T. Ottmann and I. Tomek (eds.) Educational Multimedia and Hypermedia: ED-MEDIA'94'. (pp. 468-473), Charlottesville (USA): Association for the Advancement of Computing in Education.