EFFECTIVE SOUNDS IN COMPLEX SYSTEMS: THE ARKOLA SIMULATION

William W. Gaver — Rank Xerox Cambridge EuroPARC, 61 Regent Street, Cambridge CB2 1AB, U.K.

Randall B. Smith — Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA, USA

Tim O'Shea — The Open University, Walton Hall, Milton Keynes MK7 6AA, U.K.

ABSTRACT

We designed an ecology of auditory icons which worked together to convey information about a complex, demanding simulation task, and observed users collaborating on it with and without sound. Our observations suggest that audio cues can provide useful information about processes and problems, and support the perceptual integration of a number of separate processes into one complex one. In addition, they can smooth the transition between division of labour and collaboration by providing a new dimension of reference. These results suggest that auditory icons can play a significant role in future multiprocessing and collaborative systems.

KEYWORDS: User-interface design issues; multimedia; auditory output strategies; interface metaphors; group work; observational studies

INTRODUCTION

Auditory icons are everyday sounds designed to convey information from computer systems [3, 4, 5, 7]. They are based on the observation that environmental sounds are usually experienced in terms of the attributes of source events (e.g., size, material, and force) rather than in terms of physical attributes of the sounds themselves (e.g., pitch, volume, or duration). Most current uses of sound in the interface rely on arbitrary mappings between attributes of sound and the information to be conveyed. Closer, less arbitrary mappings may be created if attributes of sound-producing events are used to represent attributes of computer events. This is the strategy used in developing auditory icons. In addition, auditory icons are often less annoying than musical messages because they can be designed to complement and extend an existing ambient auditory environment.

In proceedings of CHI, 1991 (New Orleans, LA, April 28 - May 2, 1991) ACM, New York, 1991. pp. 85 - 90.

Auditory icons are likely to be especially useful in complex systems (e.g., multiprocessing or collaborative), in which users cannot be expected to notice all visual feedback. Because users need not focus on a source to hear a sound, auditory cues are well suited to augment graphics in such systems. But the design of auditory cues for such systems is difficult because it is desirable to present a number of cues simultaneously, and some will be continuous (e.g., to indicate processing rates) while others will be discrete. Care is needed to design an ecology of sounds to work together so that each sound may be heard and understood. We have little experience with designing sounds for such environments and little user testing has been performed on such systems to date.

The research reported here explores three issues related to the use of sound in complex systems. One is the application of auditory icons in complex, demanding environments. The second is the use of auditory cues to assist collaboration. The third is the requirements for an ecology of auditory icons which work together to convey information about many simultaneous events. To explore these issues, we designed a novel simulation of a manufacturing plant and observed users collaborating on this system with and without auditory feedback.

THE ARKOLA BOTTLING PLANT SIMULATION

For this research we developed a simulated softdrink factory complex enough to provide a challenging domain for auditory icons. The simulation was implemented in the SharedARK environment [10], a collaborative version of the Alternate Reality Kit (ARK) [9]. We dubbed our simulated softdrink ARKola and named the plant accordingly.

SharedARK was created as a virtual physics laboratory to explore ideas about distance education of the future. Developed in Smalltalk, computer entities are all modelled as physical objects with masses and velocities. Users work together in a virtual space much larger than the screen, using mouse-driven hands to pick up, move, and even throw objects. SharedARK proved extremely valuable in developing our more specialized simulation as it allowed the design of virtual machines implemented as Smalltalk

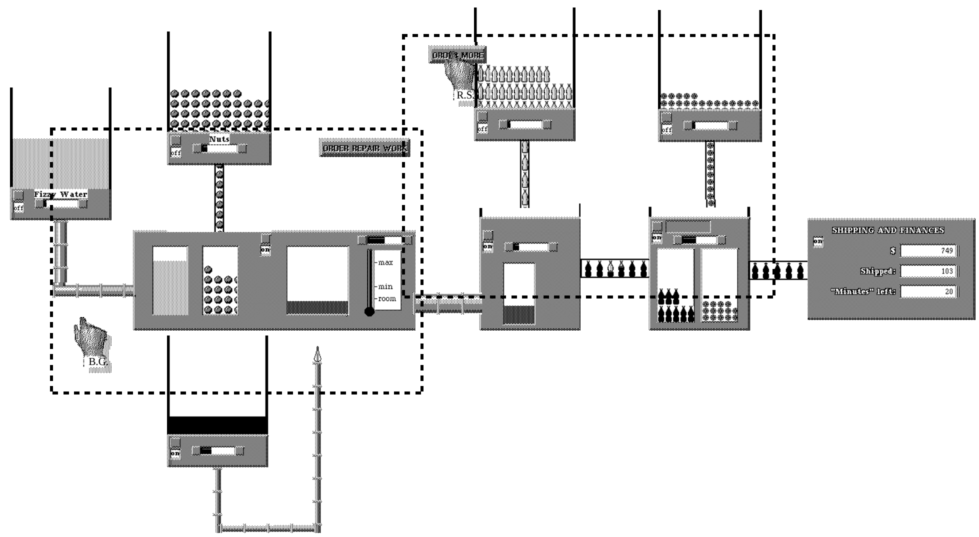

Figure 1. *The ARKola bottling plant simulation (about one fourth actual size). Rectangles show the extent of the plant each user sees at a given time.*

objects that use message passing to send their products to one another, and allowed the use of the existing interface for their presentation and control.

The ARKola factory consists of an interconnected series of nine machines, eight of which are under user control (Figure 1). There are two basic sorts of machine: outlying ingredient dispensers supply raw materials to be combined by the central line of four processing machines. The left-hand side of the plant combines fizzy water, kola nuts and heat to produce cooked cola, while the right-hand side fills and caps the bottles, and sends them to the shipping and financing machine on the far right.

Each of the machines has an on/off switch and a rate control. Users produce bottled cola and avoid wasting materials by balancing the rates of the machines, filling dispensers, and repairing occasional breakdowns generated in a predetermined pattern unknown to users. Money is made when bottles are shipped, and spent on supplies and repairs. The goal of users is to make a profit by producing as many bottles as they can as efficiently as possible.

This simulation has many attractive features for evaluating ideas about future computer interfaces. It is tractable and easily understood, yet shares much of the complexity of real systems. It provides a suitable domain for understanding how people handle many simultaneous processes, and allows exploration of audio feedback about invisible events. The task was collaborative, which allowed us to explore the ways sound might affect people working together on a complex and demanding task. Finally, it was simple to understand but relatively difficult to perform, and a great deal of fun.

DESIGNING AUDITORY ECOLOGIES

In understanding how to design an ecology of sounds for the ARKola factory, it is useful to consider the sound made by an automobile engine. On the one hand, people often experience automobile sounds as a unity, listening to whether the engine is running well or poorly. On the other hand, if something breaks the resulting noise is likely to be perceptually salient in the overall engine sound. Not only will the engine as a whole sound as if it is not running correctly, but people (especially automobile mechanics) can hear what has gone wrong.

The sounds made by automobile engines are undesigned, yet related to their sources in lawful and perceivable ways. In designing auditory icons for computer systems, we have the freedom to choose only those sounds which will be functional, to retain the close mapping between sound and source, and to shape the sounds acoustically to be discriminable. In this way, an ecology of sounds can be designed that can be heard together as an overall plant noise or attended to separately to obtain information about individual machines.

In accordance with this strategy, we designed our system so that each of the machines made sounds to indicate its status over time. The sounds were chosen to reflect the semantics of the machine. So, for example, the heater made a whooshing sound like that of a blowtorch, the bottle dispenser made the sound of clanking bottles, etc.

The rhythm of the sounds reflected the rate at which the machine was running. If a given machine ran out of supplies or broke, its sound stopped.

In addition, a number of sounds were added to indicate that materials were being wasted. For instance, a splashing sound indicated that liquid was being spilled, and the sound of crashing glass indicated that bottles had been lost. These sounds indicated what was being wasted, which didn't specify but did constrain where the wastage was occurring. We designed these sounds acoustically to contain many transients, as rapidly changing sounds convey a sense of urgency more effectively than smoother sounds [2, 11]. Finally, routine actions such as button attachments and presses were confirmed by simple, relatively discrete sounds.

Because of the complexity of the system, as many as fourteen sounds might be played at once. We attempted to design the sounds so that none would be masked (rendered inaudible) by other sounds. This involved equalizing their loudness, while making them maximally discriminable in terms of their time-varying frequency, amplitude, and timbral characteristics. In addition, we avoided continuous sounds for ongoing processes, instead using sequences of sounds repeated about once a second with silence between each burst to reduce the probability that all sounds would be playing at once. In sum, then, the auditory ecology consisted of a number of repetitive streams of environmental sounds related to the processes they conveyed and varied to reflect important parameters of these processes.

OBSERVING USERS

We watched eight pairs of subjects run the ARKola plant simulation in two one-hour sessions, one session with and one without sound. The order of the two conditions was balanced. Partners collaborated from separate rooms, communicating via a two-way audio/video connection and working on a shared version of the ARKola plant (Figure 2). Sounds were generated by a pair of MIDI-controlled digital samplers in one of the two offices and sent to the other over an in-house audio/video network [cf. 7]. We collected data from four video sources: the two cameras used for communication, and two cameras pointed at the subject's computer screens. These signals were mixed so that all four could be seen simultaneously on a single screen (see Figure 3). Audio data from each participant's microphone and from the samplers were also mixed on the data tape.

To familiarize participants with the ARKola plant, we showed them a 30-minute video tape before the first session and introduced them to the sounds before the start of the session involving sound. Subjects ran the plant for an hour each session, with a break between the first and second sessions.

THE EFFECTS OF SOUND

Figure 3 shows a screen shot of the video data we collected, while transcripts 1 and 2 show representative interactions between subjects with and without sound. These transcripts include examples of many of the features of interaction that we observed.

Listening to the Plant

Auditory cues helped participants monitor the status of ongoing processes. For instance, processing machines only made sounds when they produced an output, which helped users adjust rates of other machines. As one participant commented: "It makes me nervous when the capping machine isn't being rhythmic." In addition, diagnosing and treating problems with the plant were aided by the alarm sounds, as can be seen in Transcript 2. Subjects seemed to recognize and respond to the various wastage sounds quite reliably. Finally, subjects listened to

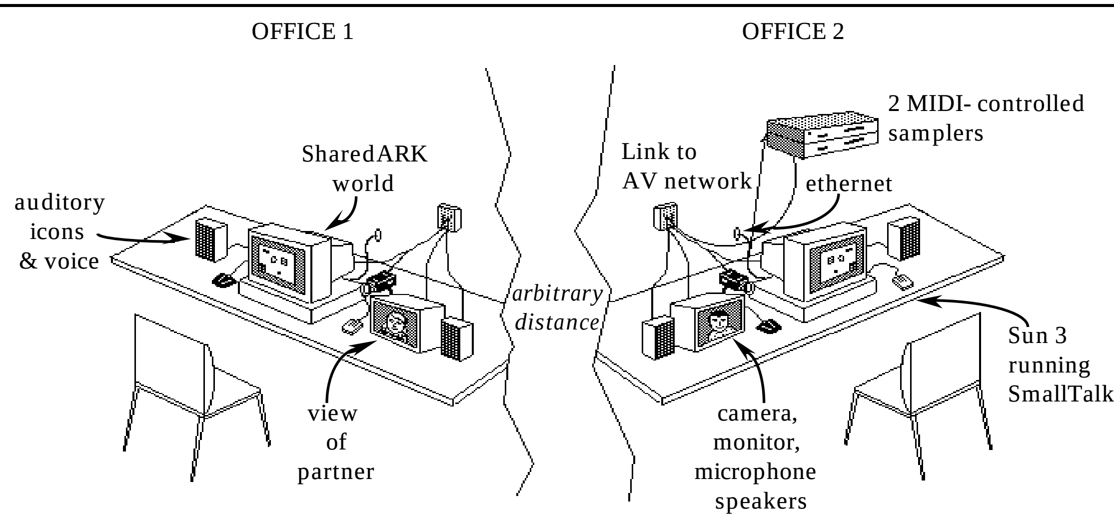

Figure 2. *Setup for the ARKola experiment. Office 1 is labelled by function; office 2 by equipment.*

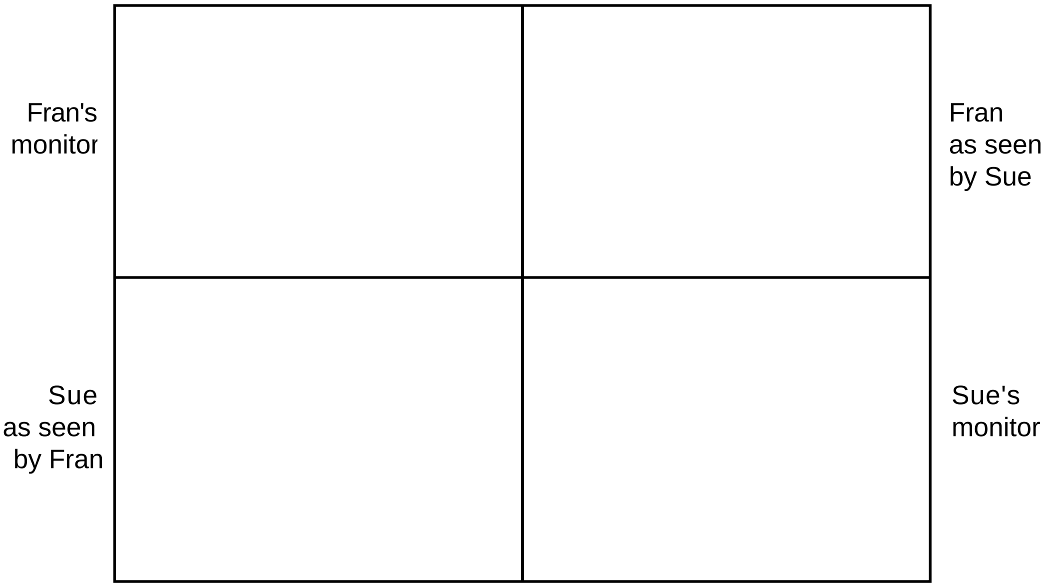

Figure 3. *A screen shot of the video data.*

Transcript 1: *Subjects: Sue and Fran session 1, without sound.*

15:28 *Sue refills nuts.*
:30 Sue: Ahh, now the maximum… we want the temperature down a bit.
:35 Sue: Is everything all right at that end?
:36 Fran: No this… I'm sure the caps aren't doing anything.
:42 Sue: Is the… you know the amount of flow through…
:47 Fran: Everything's on…
:48 *Fran changes view to include Shipping & Finances machine.*
:57 Fran: We haven't shipped anything yet and it's all going through (laughs)
:59 Sue: Where's it going then?
16:01 Fran: I don't know.
:03 Sue: Oh pooh. (laughs)
:09 Sue: The fizzy water isn't filling up in this tank so I'm going to repair it here…
:17 Sue: Can't stop these blooming nuts…
:19 Fran: Repair work, right…
:21 *Fran gets repair button, drops it on cap dispenser*
:26 *Sue moves to heater*
:27 Sue: Oops! We're using up lots of fuel here, for no purpose whatsoever.
:38 Fran: Well I've repaired the caps but it's still not working…

Transcript 2: *Subjects: Sue and Fransession 2, with sound.*

16:12 *Capper cap tank overflows; cap spill sound starts*
:14 Sue: The um… caps are coming out I think.
:16 *Fran moves to cap dispenser; turns it off.*
:23 Fran: Yeah. I've turned them off.
:35 *Bottler cola tank overflows; spill sound starts.*
:42 Sue: What's overflowing?
:43 Fran: It's… it's the… it's the whatever it is… the bottle…
:48 *Sue moves to cooker rate control.*
:50 Sue:Let me turn the pumping thing down. *(lowers cooker rate)*
…
26:16 Fran: Oh no, our money's going down instead of up.
:20 Sue: Why…What's not working?
:22 Fran: I don't know.
:28 *Sue orders fuel*
:29 Fran: Ok… no… yeah, we're going down. What are we doing?
:34 Sue: Everything sounds right, doesn't it?
:35 Fran: Yeah.
:43 Fran: Ok, it's gone up a bit.

the combined sounds of the entire plant just as people listen to the combined sounds of automobile engines (e.g., Sue 's comment in transcript 2: "Everything sounds all right, doesn't it?"). The ability for the sounds we designed to be heard separately in terms of discrete processes or as reflecting one complex process seemed quite powerful in allowing subjects to monitor individual machines and the status of the plant as a whole.

Despite their effectiveness, several problems appeared with the sounds we used. First, the urgency of the alarm sounds did not always appropriately reflect the urgency of the situations causing them. For example, the breaking bottle sound was so compelling semantically and acoustically that partners sometimes rushed to stop the sound without understanding its underlying cause or at the expense of ignoring other more serious problems. Conversely, some alarm and process sounds were not sufficiently discriminable and often went unattended by users. In particular, when a dispenser ran out of materials, its sound simply stopped. We expected this to prompt users to order more materials, but it seemed that the dying away of sounds was not salient to users. Similarly, participants who thought "everything" sounded all right were sometimes wrong – for instance, because they might not have noticed the absence of a crucial noise. This suggests that indicating the end of a process by the cessation of a sound is unlikely to be effective, particularly in environments using other sounds.

Nonetheless, most of the sounds were quite effective in allowing people to monitor events and diagnose problems. Traditional uses of sounds indicate that *something* is happening (usually something wrong), but not *what* is happening. Auditory icons seem admirably well suited for conveying semantic information about events. In addition, they allow people to maintain contact with parts of the plant they cannot see. This feature seemed to significantly affect collaboration.

Division of Labour

The nature of the simulation encouraged a division of labour among partners. This division was afforded by the loose coupling between the two parts of the plant – they were connected by only one pipe – and the fact that each half would (almost) fit on the computer screen. Participants nearly always split responsibility for the plant, with one partner controlling the cooking section, and the other controlling the bottling and capping section.

This division is well illustrated by the discussion in Transcript 1. Sue and Fran are running the plant in their first session, without sound. Sue has taken charge of cooking, while Fran is in charge of bottling and capping. Unfortunately for Fran, both the cap dispenser and the capper have broken. This means that incoming bottles are discarded, since the capper's bottle holding tank is full. Throughout this episode, she attempts to diagnose and solve the problem (E.g., "Everything's on…"), and makes repeated efforts to engage Sue in its solution ("Well I've repaired the caps…"). But though Sue pays some attention to her descriptions of the problem, Sue repeatedly goes back to her own tasks ("Oh pooh… The fizzy water isn't filling up in this tank…"). Conversely, though Sue explains the problems she is having, Fran makes no comments or suggestions about them.

This pattern of relative inattention to the other's problems changed with the addition of sound. In Transcript 2, Sue and Fran are running the plant with auditory feedback. Again, Sue is cooking while Fran is bottling. When caps overflow, Sue points it out even though it is not in her half of the plant. Similarly, when the bottler's cola tank overflows, she calls Fran's attention to it and takes an active role in solving the problem. The basic division of labour persisted in that partners split the task and became experts in their own domains. However, hearing events in unseen portions of the plant prompted users to collaborate more closely in running the plant.

Several features of auditory icons seem responsible for this change in interaction style. Sounds served as shared reference points for partners, allowing them to refer directly to events they couldn't see. In addition, sounds helped users monitor their own side of the plant, and thus freed them to attend to the other side. Finally, being able to hear processes meant that it was possible for participants to travel to the other side of the plant and continue to monitor their own side. With auditory cues, users could divide labour in the plant, maintaining different views while still being able to hear one another. If events elsewhere in the plant sounded interesting or urgent, a more directed collaboration could ensue. Adding a dimension of auditory feedback makes shared workspaces more like those in the real world, and promises to support a more casual transition between division of labour and active collaboration.

CONCLUSIONS

Our observations indicated that auditory feedback played several significant functions in participant's interaction with the system and each other. They allowed participants to diagnose problems and monitor the plant as a whole as well as individual machines. In addition, they served as shared reference points for partners, altering the nature of their collaboration.

We were impressed with the speed with which users learned the meaning and functions of this quite complex system of sounds. Semantically, the sounds worked quite well, with few subjects forgetting their meanings. The problems we did observe suggest the need for a more systematic approach to shaping the acoustics of the sounds. Psychoacoustic analyses could help improve the creation of discriminable streams of sounds [1, 8]. Similarly, research on acoustic factors leading to perceived urgency could help in appropriately indicating the importance of messages [2, 11]. Although semantic and acoustic requirements for sounds are sometimes incompatible, there is a great deal of scope for further work in this area.

The ability for an ecology of sounds to convey information about a great number of processes is clearly relevant to process control environments such as Steamer [6]. It is also relevant for many other complex systems. Subtle machine-like noises might indicate the status of the system, including background processes such as printing or networking. Many people already listen to the mechanically-produced sounds of disk drives and printers; they lose access to these sounds as systems evolve to employ remote equipment. Not only can these sounds be designed back into systems, but our work suggests greater benefits resulting from the purposeful design of such sounds. Just as the subjects we observed used auditory cues to make simulated cola efficiently, so might we all use similarly rich ecologies of sounds to accomplish our everyday work.

ACKNOWLEDGEMENTS

We thank Victoria Bellotti, Allan MacLean, and Tom Moran for their comments on this paper, and our subjects for their participation in the experiment.

REFERENCES

1. Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, Mass., MIT Press.

2. Edworthy, J., Loxley, S., Geelhoed, E., & Dennis, I. (1989). The perceived urgency of auditory warnings. *Proceedings of the Institute of Acoustics 1989 Spring Conference. 11*, (5), 73 - 80.

3. Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction. 2*, 167 - 177.

4. Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. *Human-Computer Interaction.* 4 (1).

5. Gaver, W. W., & Smith, R. B. (1990). Auditory icons in large-scale collaborative environments. *Human-Computer Interaction – Interact'90*. D. Diaper et al. (eds.) Elsevier, North-Holland.

6. Hollan, J. D., Hutchins, E. L., & Weitzman, L. (1984, Summer). STEAMER: An interactive inspectable simulation-based training system. *AI Magazine*, pp. 15 - 27.

7. Moran, T. P., & Anderson, R. J. (1990). The workaday world as a paradigm for CSCW design. *Proceedings of CSCW 90* (Los Angeles, U.S., October 1990).

8. Patterson, R. D. (1989). Guidelines of the design of auditory warning sounds. *Proceedings of the Institute of Acoustics 1989 Spring Conference. 11* (5), 17 - 24.

9. Smith, R. B. (1987). The Alternate Reality Kit: an example of the tension between literalism and magic. *Proceedings of CHI + GI 1987* (Toronto, Canada, April 5 - 9, 1987) ACM, New York, 61 - 67.

10. Smith, R. B. (1989). A prototype futuristic technology for distance education. *Proceedings of the NATO Advanced Workshop on New Directions in Educational Technology*. (Nov. 10 - 13, 1988, Cranfield, U.K.)

11. Swift, C. G., Flindell, I. H., & Rice, C. G. (1989). Annoyance and impulsivity judgments of environmental noises. *Proceedings of the Institute of Acoustics 1989 Spring Conference. 11* (5), 551 - 555.