Sound Support For Collaboration

William W. Gaver

Rank Xerox EuroPARC, England

Shared work often involves fluid transitions between relatively focussed collaboration, division of labour, general awareness and serendipitous communication. This leads to a tension in the design of software systems meant to support shared work: focussed collaboration implies the need to coordinate people's views of work objects, while division of labour requires individual control over views. A similar tension exists in the office environment as well: group engagement in the workplace depends on a shared context, but individual work is facilitated by privacy and freedom of action. Auditory cues have the potential to reduce these tensions because graphics and sound can provide two independent ways to present and obtain information. I illustrate the potential of sound in collaborative systems with observations drawn from two systems: the ARKola simulation, which explores the effects of sound on collaboration within a workstation environment; and EAR, in which auditory cues are used to increase general awareness of events and encourage group engagement within the workplace itself. These examples suggest useful functions sound can play in collaborative systems.

Introduction

The shift from computer systems that support a single user working alone to those supporting a group of users working together is a profound one. It leads to a consideration of the ways people work together in the everyday world and possible ways to extend and support their interactions. Perhaps more importantly, it suggests that the unique capabilities of computers should be embedded more firmly in ordinary work practises, so that the distinctions between the world of the computer and the workaday world are blurred (Moran & Anderson, 1990).

Gaver, W.W. (1991). Sound support for collaboration. Proceedings of the Second European Conference on Computer-Supported Collaborative Work.

Developments in collaborative systems are promising, but if traditional models of human computer interaction seem to assume that we work in isolation, the new model sometimes seems one of people spending the totality of their working lives in meetings. To date, most software systems designed to support shared work seem aimed at supporting relatively intensive periods of collaboration – for instance, in meetings (Mantei, 1988), creating structured outlines (Ellis et al., 1988), or simultaneously developing documents (CSMIL, 1989).

But just as most people don't work alone at all times, nor do they always work together. Often people are merely aware of each other – aware of others' presence, perhaps their activities and progress. Occasionally people meet randomly in the course of day to day work, and these meetings are serendipitously fruitful, as when casual conversation leads to some question being answered or a longer term collaboration being started. And even when collaborating, people often divide their labour, meeting one another to share results and plan the future. Only occasionally do we actually join and work together closely on the same task.

People shift from working alone to working together, even when joined on a shared task. Building systems that support these transitions is important, if difficult. One promising approach is to embed collaborative software in a larger system of audio and video interconnectivity that allows people to be virtually co-present even if not working closely with one another (e.g., Buxton & Moran, 1990; Root, 1988; Goodman & Abel, 1987). Such systems have had some success, but it also seems important for such transitions to be supported by software systems themselves.

In this paper, I discuss the potential for auditory cues to support relatively casual and serendipitous forms of collaboration, both in software and office environments. First, I explore the movement between awareness and focussed collaboration, and discuss the reasons auditory cues seem appealing for support of smooth transitions in the degree of engagement on a common task. The potential of auditory cues is illustrated with examples from two systems that use sound to support collaboration. The first example comes from the ARKola bottling plant simulation, which explores the effects of auditory cues on a collaborative task in a workstation environment. The second system, called EAR (for Environmental Audio Reminders), is a system in which sound helps users maintain awareness of one another and events within the workplace itself. These two examples complement one another in focussing on the effects of auditory cues on collaboration in the workstation environment and the more general office environment; together they point the way towards many possible future developments.

Moving among ways of working

Figure 1 is a simple representation of the complex process of working together. Although simplistic, it provides a useful orientation to the extremes of the experience. Four major landmarks are indicated here. Underlying all is *general* *awareness*. This is a pervasive experience, one of simply knowing who is around and something about what they are doing: that they are busy or free, meeting or alone, receptive to communication or not. Awareness is necessary for all collaborative work, but the degree to which its focus is shared varies. An intense sharing of awareness characterizes *focussed collaboration* – those occasions in which people work closely together on a shared goal. Less is needed for *division of labour*, that common work practise in which a shared goal is divided and component tasks addressed separately. Finally, more casual awareness can lead to *serendipitous communication*, in which people realize the potential for productive work through chance encounters.

People move among these ways of working together along many trajectories. Simple awareness may lead to serendipitous communication, which in turn may lead to division of labour or focussed collaboration. Alternatively, a period of focussed collaboration may be followed by a division of labour. All of these forms of working together are likely to be important at one time or another in a shared project; supporting fluid movements among them is an important goal for collaborative software.

Yet the design of systems with the flexibility necessary to support many styles of shared work is not an easy task. One problem seems to be the tension between the need to maintain a common focus for collaborators and the desire to allow individuals freedom to work on their own. Bellotti et al. (1991) make this tension explicit in a Design Rationale based around studies of a shared editor (cf. MacLean et al., 1989). Two of the criteria they identify as pervasive in the choice of design

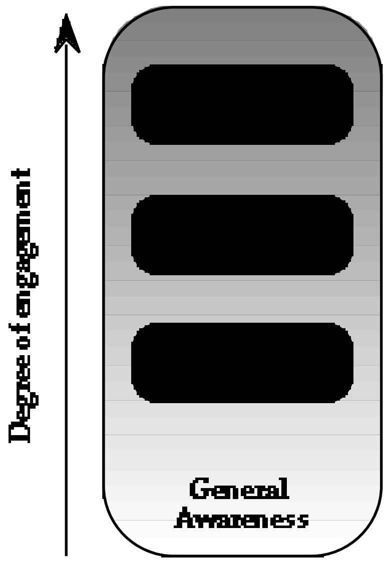

Figure 1: Shared work involves fluid transitions among focussed collaboration, division of labour, serendipitous communication, and general awareness (which supports them all).

options are the seemingly contradictory ones of "maintaining shared work focus" and "allowing individual work."

In the design of shared software systems, the tension between shared and individual work is reflected in issues concerning the degree of control over work objects afforded users. Individual work is supported by giving people complete control over their view of a work object: over its screen placement, the parts of it made visible, their appearance, etc. But shared focus is supported by *reducing* individual control over their view. From this perspective, focussed collaboration is most likely to occur when all participants can be assumed to be viewing the same thing. Although enforcing an identical focus on a given task may be helpful for supporting focussed collaboration, it is likely to hamper the smooth flow to other, less close forms of shared work (Bellotti et al., 1991).

Similar issues arise in offices, where the shared contexts necessary for group engagement compete with the privacy needed to concentrate on individual work – it is difficult to get work done when constantly in meetings about work. Providing ubiquitous audio-video interconnectivity may encourage awareness, but one must monitor a video screen at the expense of attention to one's work. Using video windows on a workstation is only a partial solution, since they must vie for valuable screen real-estate with other graphical tools.

In sum, systems which seek to support both shared work and individual flexibility suffer from the need to compete for control over the same display resources and limited visual attention. Clearly these issues can be dealt with by increasing the size and number of displays and relying on the time-honored panacea of social control. In this paper, however, I suggest that sound can provide a valuable alternative to vision as a means of providing the contextual information that allows free movement among more and less intense forms of collaboration.

Auditory icons and collaborative work

There are a number of reasons to think that sound has the potential to complement visual displays in supporting the transitions between focussed collaboration and more casual and separate forms of shared work. Primary among these reasons is hearing's status as a distance sense secondary only to vision. By distance sense, I mean that we are able to listen to information about events at a distance. Just as we can see a tree fall from far away, so can we hear it. We hope, on the other hand, neither to feel or taste the falling tree; and though we may smell it the experience is not likely to provide us with much useful information.

Because we can listen to as well as look at distant events, we can divide information about computer events between the two senses. On the one hand, we may provide redundant information about an event, so that we can both see and hear it. More interesting, we can disassociate the two, so that we may hear what we don't see.

Hearing also complements vision in that listening to an event does not necessarily interfere with the maintenance of a visual focus on another event. As I write this, for instance, I might hear a colleague walk by my office. The sounds of footsteps, doors opening, etc., provide information about what is going on around me, but I can nonetheless maintain my focus on my work. This should carry over quite well to collaborative systems, so that individual control can be granted users while sufficient cues as to the activities and whereabouts of others are still available. By splitting information about a shared workspace between sound and vision, we may reduce the tension between the desire to maintain a shared focus and that of allowing individual work.

Of course, no matter how attractive sound may be as a medium, it must be able to convey relatively complex information about events if it is to be useful. Clearly a collaborative system relying on the beeps and buzzes currently used in computers to increase awareness of colleague's activities would entail too high a cognitive overhead to provide valuable support to users (not to mention the irritation it would cause). It is not only necessary that sound complement vision, but that it provide information in subtle and intuitively obvious ways.

I have been developing a strategy for using sound to convey complex information that is based on the ways people listen to events in the everyday world (Gaver, 1986). From this perspective, we listen not to sounds and their attributes (such as pitch, loudness and timbre) but rather to events and theirs (e.g., footsteps, force and size). *Everyday listening* refers to the experience of listening to events. Taking this experience of listening as primary allows the development of a framework for analyzing and manipulating sounds that is based on attributes of events rather than the parameters of sound *per se*. These attributes, in turn, may be mapped to attributes of computer events, giving rise to *auditory icons.* Auditory icons are environmental sounds (like taps, scrapes, etc.) designed to convey information by analogy with everyday sound-producing events.

Auditory icons have several appealing qualities as a method of providing feedback about events. First, sound as a medium is a valuable way to provide information that is not constrained to a single location (e.g., I can hear a sound without facing my computer monitor). Second, non-speech audio is often less distracting, less susceptible to masking, and more efficient than is speech. Third, everyday sounds can often be mapped more closely to the events they are meant to represent than can musical sounds. Finally, auditory icons can be designed to present information in an almost subliminal way – just as we are likely to get a great deal of information without conscious attention from the sounds of colleagues working, so can auditory icons convey a great deal of information without being overly distracting.

Experience with systems employing auditory icons has suggested that such cues can be useful for individual work (Gaver, 1989). In particular, sound can convey information about events and objects that is difficult to convey visually – for instance, about the timing of events or the nature of interactions – as well as information that is inconvenient to present and obtain visually, for instance about the progress of relatively long lasting processes. Finally, informal experience with sound in a large-scale, collaborative system called SoundShark (Gaver & Smith, 1990) suggests that sound can support general awareness of collaborators' whereabouts and activities.

What I am suggesting, then, is that a smooth flow from focussed collaboration to division of labour can be facilitated by using well-designed auditory icons to increase awareness of activities and events. In the next two sections, I expand and support this notion by detailing experience with two collaborative systems which employ auditory icons. The first is the ARKola bottling plant simulation, a system in which sound provides cues designed to aid users collaborating in a workstation environment. The second is EAR, a system that uses designed audio cues to support awareness of events and activities within the entire work environment.

The ARKola bottling plant simulation

The ARKola bottling plant is a simulation designed expressly to explore the functions of auditory cues in complex, collaborative software systems. The simulation was developed to serve as a domain for testing that would satisfy a number of constraints:

- We hypothesized that sounds would aid in monitoring multiprocessing systems, so many simultaneous processes should be involved in the task.
- Sound should enable people to track hidden or invisible events, so the task domain should be too big to entirely fit the computer screen.
- Auditory cues are likely to be most evidently useful when tasks are demanding, so we wanted a task that was simple to understand yet difficult to perform.
- We expected sound to affect collaboration, and so wanted a task that would encourage shared work.
- Finally, we wanted a task that would seem natural and engaging for participants, so they would not be bored or confused during our studies.

The ARKola simulation seemed to fulfill these requirements quite well. We stress, however, that though this simulation may seem more representative of video games or process control tasks than of traditional workstation domains, we believe it shares many features with – and thus our results are relevant to – more traditional domains. Although we were interested in testing several functions for auditory icons within this environment, for the purposes of this paper I focus primarily on aspects directly relevant for collaborative work (for a more complete description of this work, see Gaver et al., 1991).

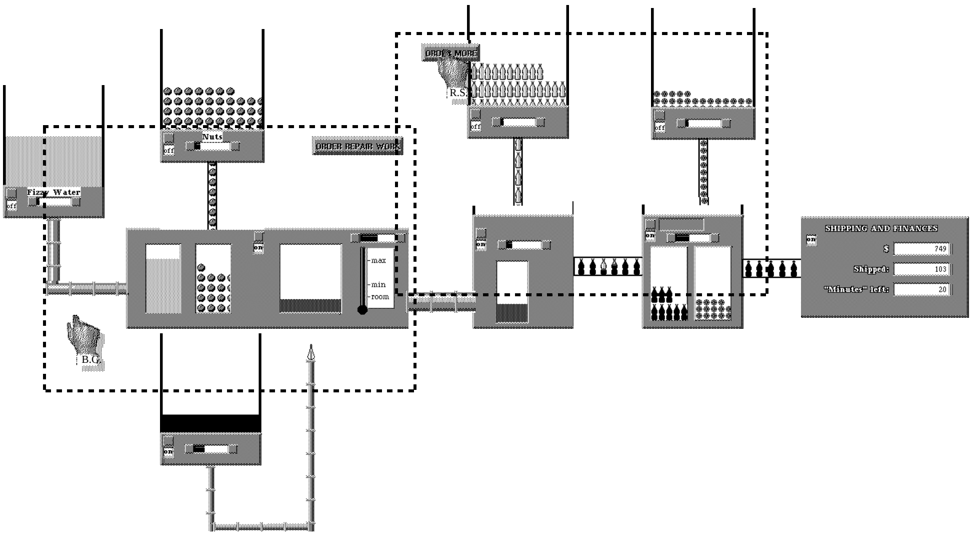

Figure 2: The ARKola bottling plant simulation. Nine machines mix, cook, bottle, cap, and count bottles of simulated cola. Mouse-driven hands are used to move and press buttons, control machines, etc. Dotted rectangles show the approximate extent of the view each user could have of the plant. (This figure is approximately one-fifth actual size.)

The plant, shown in Figure 2, consists of a virtual assembly line for producing a simulated softdrink. Users control the plant using mouse-driven "hands" to activate machine controls and to move and activate "buttons" which order new supplies or repair broken machines. Completed bottles of cola add funds to a virtual "bank account" at the end of the line; buying supplies or repairs deplete funds. The goal of participants, then, was to make money by producing as much cola as possible as efficiently as possible.

The simulation was implemented in SharedARK, a collaborative version of the Alternate Reality Kit (Smith, 1987); thus the simulated softdrink was called ARKola and the plant named accordingly. SharedARK, a fascinating environment in its own right, was used here as a foundation for developing the visual appearance and actions of the plant and participants' interactions with it.

Feedback about the status of the plant was provided by visual and auditory cues. Supplies could be heard as they moved along: cooking cola burbled, the capping machine clanged, and wasted supplies crashed and spilled audibly. Although some attempt was made to equate the information presented audibly with that displayed graphically, the purpose of the experiment was not to compare the two media in terms of effectiveness, but rather to understand their different characters.

The bottling plant was designed to be too large to fit on a computer screen, so each participant could only view part of the plant at a given time. However, participants could move their view by "sliding" their screen over the plant. Thus people could coordinate their views to work with a shared focus, or use separate views and divide their labour.

Observing collaboration on the plant

We observed eight pairs of people using the system for two one-hour sessions apiece; one session with and one without auditory feedback. Half the participants had auditory feedback on their first sessions and half did not. Partners worked on the system from different offices in the building, working together in the "same" factory shown on different workstations and communicating via a two-way audio and video link. Figure 3 shows the experimental set-up for the two offices.

We collected video-taped data upon which we based our observations of plant usage both from the subjects' audio-video links and from cameras pointing at each of their screens. Our observations are informal, relying mainly on occasions when participants explicitly referred to the sounds. We were able to cull a number of suggestive examples of the use of sounds. We take our data, then, as providing hypotheses for further testing and exploration.

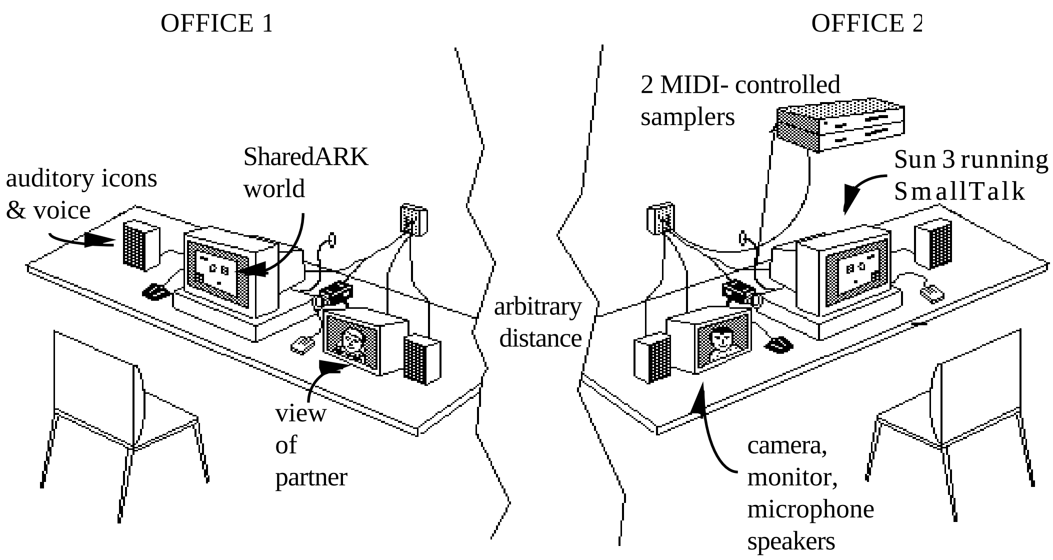

Figure 3. Setup for the ARKola experiment. Subjects worked in separate offices, collaborating on the ARKola simulation and communicating via an audio-video link. Data was collected from their camera and from cameras pointing at the computer screens.

Collaboration in the ARKola Simulation

We were struck with the great degree to which participants divided the labour of running the system in this study. We had not expected this, but our observations indicated that division of labour was encouraged by the design of the simulation. The plant divides rather neatly into two halves, with the four machines on the left (which produce cola) connected to the five on the right (which bottle and cap it) by only one pipe. In addition, the operation of the cooking half did not depend on bottling at all, while incoming cola was buffered by the bottling machine, reducing time dependency on the cooking half. Because the two sides were relatively independent, then, and because there was only one connection between them, each could be run without much care for the other – though of course successful performance on the task itself required that both sides be well run.

The tendency for participants to divide the task was made apparent by the large amounts of time that they spent using different views on the system. After an initial period during which the partners would usually wander over the plant together in order to orient themselves to the machines, they almost always separated and seldom shared views again. (For instance, participants M. and H. made this explicit. M: *"Maybe this is a good strategy, actually, to look after half of the world each..."* H: *"Yes, then we can... keep an eye on machines and see them break straight away."*) This division of labour was also made evident in their conversations. Although each would comment to the other about events and progress on their respective sides of the plant, longer conversations in which the two would collaborate on solving a problem were relatively rare.

The addition of auditory cues seemed to change this pattern of division of labour to a noticeable extent. Although subjects still maintained separate views to a great degree, their conversations seemed to reflect a greater degree of concern for events on their partner's side. (For instance, in one tape E is working on the cooker half and P on the bottling half. P remarks on a sound made by a machine on the other side of the plant: *"Isn't that the fizzy water that's leaking?"* E: *"I don't think it's leaking... I think it's just going into the tank."* Caps start spilling on P's side. P: *"Ok, I'm losing, uh..."* E: *"That's the caps."* P: *"caps..."* P turns off cap dispenser.) While joint problem-solving was relatively rare without sound, it became common with auditory feedback.

The ability for both partners to hear events seems to be the key to sounds' effect on their collaboration in this task. Running the ARKola simulation was relatively demanding, requiring constant attention to the state of supply hoppers and the flow of materials through the plant. Leaving one's area of responsibility was risky in that some disaster was liable to occur; without auditory feedback this would go unnoticed until one's return. Because partners could not see each others area of the plant, joint problem-solving required verbal descriptions of plant status and made problem solving much more difficult for the distant partner. Each participant tended to focus on his or her own responsibilities, without the possibility of direct awareness of other events.

Auditory feedback allowed users to be aware of parts of the plant that were not visible on the screen or at the focus of their visual attention. Thus participants could refer directly to sounds from their partner's half of the plant and hear problems occurring in areas on which they were not focussing. (For example, when bottles started breaking on T's side of the plant, his partner, S, said: *"Bottles are breaking!"* T: *"Where?"* S: *"I don't know, but they're breaking..."*) Being able to hear the status of the plant also reduced the risk of venturing to other areas of the plant. If problems did occur during one's absence, they were likely to be heard. In providing a new dimension of reference for partners running the plant, auditory cues seemed to ease the transition between division of labour and collaboration in this system.

Of course, the sounds we used were not without their problems. Care was needed to ensure that the auditory feedback was loud enough to be heard without preventing conversation, for example – though this is not a difficult task, it is a crucial one. In addition, designing the sounds to work together so that all could be heard was quite demanding (see Gaver et al. 1990 for a description of our approach to this problem). Finally, some of the sounds were more effective than others. Most notably, when a supply hopper ran out of supplies its sound simply stopped. We had expected that participants would notice the cessation of sound and refill the hopper; instead the sound's absence often went unnoticed. Nonetheless, the majority of sounds seemed informative and useful to subjects.

In sum, the auditory feedback used in this system had important effects on participants' collaboration. Sound provided a new dimension of reference for subjects. By increasing ways to maintain awareness it smoothed the transition between division of labour and focussed collaboration. Being able to hear the status of offscreen machines allowed a dissociation of focussed visual attention and more general awareness, so that each participant could have an area of primary responsibility and still join together to solve problems.

It is important to stress that we expect these findings to be relevant to a broad range of shared software, not just the sort of process control simulation described here. As systems become more powerful, they are increasingly likely to demand the scheduling and control of simultaneous tasks which are often hidden or invisible – and collaborative as well. The ARKola simulation was designed as it was precisely to embody these features in a self-motivating task domain, so that our results would be broadly relevant.

Our observations of the ARKola simulation in use are indicative of the potential for auditory cues in collaborative software systems. Such cues can also support awareness of events and activities beyond the computer in the encompassing workplace. I explore these possibilities in the next section.

Ambient audio in the workplace

A great many collaborative activities take place in the office environment, ranging from relatively formal meetings to more casual encounters. Just as there is a tension in collaborative software systems between enforcing a shared focus and allowing individual activities, so are there tensions in the workplace between encouraging group engagement and providing for individual work. As with collaborative software, group engagement in the workplace depends on a shared context – meeting rooms, open spaces, and established office hours. But individual work is facilitated by individual control over the environment – private offices, work at home, or work during off-hours.

In the everyday world, this tension is mitigated to some degree by the naturally-occurring auditory environment. We often listen to ambient sounds in the workplace in order to maintain awareness of our colleagues' activities. As I write this, for instance, I can hear automobiles and buses pass by on the street below, people walking by outside my office, and the sudden roar of the copier machine being used. As with collaborative software, these sounds may provide the sorts of awareness useful for moving in and out of close collaboration. For example, hearing Paul enter his office next door may prompt me to ask him about some project of mutual interest. Hearing the murmur of voices from outside my office may encourage me to join in an informal discussion with my colleagues. Hearing nearby events in the building can support casual awareness of others or indicate ongoing meetings, whether serendipitous or formal.

Hearing events in the workplace can draw us into them; but in large buildings many will go unheard. In addition, many potentially relevant events don’t make informative sounds. For instance, hearing Paul leave his office may tell me he is unavailable, but not whether he is going to a meeting, to fetch some coffee, or to the pub. And of course, naturally-occurring sounds can be irritating, as sounds of the rush of traffic, the roar of the copier, and the blare of Paul's stereo often are. Such sounds are annoying because they are not informative or relevant: Noise is uninformative sound. In general, the ambient audio environment of the workplace can be useful, easing the tension between group and individual work. But sound can also pose problems: not all events may be heard, some important events may not make sounds, and the sounds events do make may be annoying.

EAR: Environmental Audio Reminders

For the past year and more, we have been using a system at EuroPARC which allows us to design informative ambient audio environments in our workplace. Called EAR, for Environmental Audio Reminders, this system triggers short, unobtrusive audio cues which are transmitted to offices around the building in order to inform people about ongoing events or to remind them about upcoming ones. Using this system, we can smoothly expand the naturally-occurring office ambience so that we can hear events out of earshot, and events which don’t ordinarily make sounds. This work can be seen as moving auditory icons out of the workstation and into the world, so that the working environment itself becomes the interface. From this perspective, the strategy guiding the use of sound to facilitate collaboration in workstation environments can be applied to the overall work environment as well.

The EAR system relies on two interesting features of EuroPARC's environment. The first is a data-base of events called Khronika (Lövstrand, 1991) which allows a wide range of events to be browsed, edited and indicated by various cues. Khronika controls the generation of audio cues which are routed to speakers in particular rooms by the second system, a computer-controlled audio-video network (Buxton & Moran, 1989). The net result of this environment is that events generate designed audio cues that can be heard remotely.

As with the design of auditory icons for workstation environments, two design constraints are important in shaping the auditory cues used in EAR. First, the sounds must be semantically related to the messages they are meant to convey. This is achieved by using sampled environmental sounds that are either causally or metaphorically related to their referents. The second constraint is that they be acoustically shaped to avoid distraction and annoyance. Our strategy for creating unobtrusive sounds has been guided by work on designing sets of auditory alert sounds of appropriate perceived urgency (Patterson, 1989). For instance, most of the sounds we are designing have relatively slow onsets, which means they do not startle or distract listeners but instead slowly emerge from the natural auditory ambience of the office. In general, we try to maintain a balance between designing auditory cues that have clearly recognizable semantic content and designing them to be acoustically appropriate.

EAR in action

EAR is used to play audio cues which support casual awareness of one another, indicate opportunities for casual (and perhaps serendipitous) communication, and inform us about more focussed and formal events in our working environment. For instance, meetings are signalled by the sound of murmuring voices slowly growing in number, ended by the sound of a gavel. The sound interrupts individual work discreetly, reminding the listener about a prior engagement to join with other members of the lab. We view teatime, on the other hand, as an opportunity for informal communication. Each afternoon people in the building are invited to take tea by the sound of boiling water, followed by sounds of pouring water and spoons stirring in teacups. This sound serves as a more gentle reminder to those of us concentrating on our work that we might want to join our colleagues. Finally, sounds have evolved to indicate even very informal meetings. For instance, in the evening one of us is likely to trigger the pub-call, which plays the sounds of a pint being poured in a background of people talking and laughing.

These sounds serve as unobtrusive yet effective announcements of events in the workplace. They don't interrupt ongoing work, and can easily be ignored (though meeting sounds are likely to be heeded). Because they are stereotypical versions of the sorts of sounds we might hear around the building every day, the auditory cues used in EAR provide an effective and intuitive way to call people together and keep them informed of events around the building.

The EAR system also uses a number of auditory cues to indicate events in the electronic environment. For instance, the arrival of email can be accompanied by the sound of several pieces of paper falling on a surface, like letters falling through a mail slot. Other auditory cues are valuable in maintaining our awareness of the status of our audio-video network. This network allows people to connect their monitors to cameras around the building to gain a sense of "virtual co-presence" with distant colleagues. Because there is no visual indication when somebody accesses the signal from a camera (and video symmetry is not enforced), a pervasive sense of monitoring might be expected to result. But the EAR system allows audio feedback about connections, so that when somebody connects to my camera I hear the sound of a door creaking open; when they disconnect I hear the door shut. These simple audio cues provide invaluable feedback about the state of the audio-video network and seem to bolster feelings of privacy control to a significant degree. In addition, they can serve to tell us about a wider context of activities than is revealed by the network alone. For instance, auditory cues are used to distinguish the purpose of an audio-video connection: Different sounds indicate "vphone" calls; casual, one-way glances; and camera accesses by our framegrabber service.

Many of the sounds we use in EAR may seem frivolous because they are cartoon-like stereotypes of naturally-occuring sounds. But it is precisely *because* they are stereotyped sounds that they are effective. Using sounds that mimic those made by actual events means that the mapping between the information to be conveyed and the sound used to represent it can be quite close, and thus easy to learn and remember. While the sounds we use must be introduced to new users, they are quickly understood and seldom forgotten. It seems unlikely that more "serious" sounds – such as electronic beeps or sequences of tones – would be as effective at providing information in an intuitive and subtle way.

Like the sounds in the ARKola simulation, the cues about our electronic environment indicate computer events without demanding visual attention. But because the primary purpose is to provide cues about events in the workplace, the system has the further effect of bringing the two environments closer. The workstation is no longer the sole source of information about the electronic environment; instead electronic events are made an integral part of the general environment.

EAR is an installed system, constantly evolving to reflect our current needs and opinions about the auditory cues. Thus we have taken a strategy of "evaluation by use," in which cues which do not seem useful or which are annoying are discarded or redesigned. Generally this evolution has involved the introduction of subtle variations between cues. For instance, soon after the door-opening sound was introduced to indicate camera accessing, new sounds appeared which differentiate between short connections made by colleagues and connections made by an application which digitizes images and makes them available to colleagues overseas.

In sum, the auditory cues used in the EAR system can be unobtrusive, informative, and valuable. They serve to indicate events in the same way that they might be heard in everyday life, with the added advantage that the events cued are chosen by users. They allow us to hear distant events, or events that don't naturally produce informative noises, helping to blur the distinction between the electronic and physical environments. Perhaps most importantly, by informing us about ongoing events in the building they help to ease the transition between working alone and working together.

Discussion

The ARKola simulation and EAR system complement one another as examples of the use of auditory cues in collaborative systems. Where the ARKola simulation explored the design of auditory cues that support collaboration within the workstation environment, the EAR system demonstrates that similar principles can guide the design of useful auditory cues in the more general working environment as well. In ARKola, auditory cues were crucial sources of information, whereas in the EAR system sounds generally support a relatively unconscious awareness of ongoing events.

But though the two systems are different in many ways, parallels can be drawn in the functions auditory cues perform in each. In both the ARKola simulation and the EAR system, auditory cues make use of sound as a new medium for increasing awareness of events and activities which are not visually available. The effect of this new dimension of reference seems to be that users can simultaneously maintain visual attention on a potentially shared focus of work while remaining aware of a wider context of interest. This ability, in turn, seems to lead to smoother transitions between different ways of sharing work.

The functions auditory cues play in the ARKola simulation and the EAR system should be broadly applicable to a number of CSCW systems. The tension between maintaining a shared focus and allowing individual control over work seems common in the class of collaborative tools that allow synchronous editing of objects (Bellotti et al., 1991). Our observations of the ARKola simulation suggest that this tension may be reduced by exploiting sound as an alternative medium for presenting and receiving information. So, for instance, users of a shared document editor might hear their partners' editing operations even when such activities are offscreen. Such sounds could be useful in coordinating activities ("...it sounds like you're making major changes up there – should I hold off on this section?").

Similarly, experiences with EAR suggest that using auditory cues to communicate contextual information in the workplace itself can facilitate the flow of engagement among colleagues. For example, just as EAR allows us to hear activities in distant parts of our building, so might users of systems supporting virtual co-presence hear activities in distant environments. Such sounds could provide a natural means for indicating potentials for casual or focussed engagement by conveying contextual information which might otherwise be lost.

I have shown in this paper some of the functions sound can perform in collaborative systems. But it should be stressed that our work on the use of auditory cues to facilitate collaboration has only just begun. Both the ARKola simulation and EAR are suggestive, but neither is definitive; the potential of sound as an intuitive, unobtrusive medium for communication promises to be much richer than either of these applications can show.

Acknowledgements

The ARKola study was done in collaboration with Randall Smith and Tim O'Shea; the EAR system with Lennart Lövstrand. I thank Allan MacLean for his many insightful comments on this work, and give particular thanks to Victoria Bellotti and Paul Dourish for many helpful discussions about the nature of sharing.

References

Bellotti, V., Dourish, P., & MacLean, A. (1991). From users themes to designers DReams: Developing a design space for shared interactive technologies. EuroPARC/AMODEUS Working Paper RP6-WP7.

Buxton, B., & Moran, T. (1990). EuroPARC's Integrated Interactive Intermedia Facility (IIIF): Early experiences. *Proceedings of the IFIP WG8.4 Conference on Multi-user Interfaces and Applications*, Heraklion, Crete, September.

CSMIL (1989). ShrEdit: A multi-user shared text editor: Users manual. Cognitive Science and Machine Intelligence Laboratory, The University of Michigan.

Ellis, C., Gibbs, S., & Rein, G. (1988). Design and use of a group editor. MCC Technical Report Number STP-263-88.

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

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

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.

Gaver, W. W., Smith, R. B., & O'Shea, T. (1991). Effective sounds in complex systems: The ARKola simulation. *Proceedings of CHI 1991,* New Orleans, April 28 - May 2, 1991, ACM, New York.

Goodman, G., & Abel, M. (1987). Communication and collaboration: Facilitating cooperative work through communication. *Office: Technology and People, 3* (2), 129 - 146.

Lövstrand, L. (1991). Being selectively aware with the Khronika system. In the Proceedings of ECSCW'91, Amsterdam, The Netherlands.

MacLean, A., Young, R., & Moran, T. (1989). Design Rationale: The argument behind the artifact. *Proceedings of CHI'89: Human Factors in Computing Systems*, April 30 - May 4, Austin, Texas, 247 - 252. New York: ACM.

Mantei, M. (1988). Capturing the Capture Lab concepts: A case study in the design of computer supported meeting environments. *Proceedings of the Conference on Computer-Supported Collaborative Work.* Portland, Ore., September 1988, 257 - 270.

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).

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.

Root, R. (1988). Design of a multi-media vehicle for social browsing. *Proceedings of the Conference on Computer-Supported Collaborative Work.* Portland, Ore., September 1988, 25 - 38.

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.)

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.