Tangible Progress: Less Is More In Somewire Audio Spaces

Andrew Singer, Debby Hindus, Lisa Stifelman*, and Sean White†
Interval Research Corporation
1801 Page Mill Road, Building C
Palo Alto, CA 94304
+ 1.650.424.0722
{singer, hindus}@interval.com

ABSTRACT
We developed four widely different interfaces for users of Somewire, a prototype audio-only media space. We informally studied users' experiences with the two screen-based interfaces. We prototyped a non-screen-based interface as an example of a novel tangible interface for a communication system. We explored the conflict between privacy and simplicity of representation, and identified two unresolved topics: the role of audio quality and the prospects for scaling audio spaces beyond a single workgroup. Finally, we formulated a set of design guidelines for control and representation in audio spaces, as follows: GUIs are not well-suited to audio spaces, users do not require control over localization or other audio attributes, and awareness of other users' presence is desirable.

KEYWORDS
Audio, speech interactions, mediated communication, computer-mediated communication, CMC, user interfaces, representations, media space, audio space, audio-only, tangible interactions, active objects, design guidelines.

INTRODUCTION
Although the technology for Internet telephony is developing rapidly, most of its interfaces and the interactions that they provide are currently tied to the physical model of the telephone. As the technology improves and the underlying systems offer the full flexibility of digital audio, the possible interactions expand and new kinds of group communication systems become practical.

One promising example of these new systems is the audio space. An *audio space* is an audio communication system for a group, the members of which are in disparate physical locations; the audio space creates the auditory illusion for each member that its users share a common acoustic space.

To explore such audio spaces, we developed the **Somewire** system. Somewire allowed us to look at design issues and the workplace influence of a high-quality, audio-only group communication system. One or more versions of Somewire were in continuous use by 8 to 16 people within Interval Research Corp. from 1993 to 1995.

A previous paper presented the findings from a Somewire-related field study [1]. This study examined usage, conversational content, and social norms within a single workgroup. The findings clearly show the value and utility of the Somewire audio space. Another significant finding of this study is that audio spaces can lead to social spaces.

The current work looks at the user interface aspects of Somewire, rather than at the social aspects. We built four interfaces to Somewire; each embodies a different approach to the central matters of representation and interaction in such a communication space. We conducted user studies and evolved the interfaces based on the results. Two interfaces are screen-based GUIs, one is a device interface with no software controls, and one is a tangible user interface that used physical objects. This first application of a tangible user interface to a media space is by far the most novel and engaging of the four; we describe it in detail.

In this paper, we first review prior work that informed the design of our audio-only media space. We describe Somewire briefly, then examine the four interfaces, emphasizing representation and interaction. We discuss what we learned from the user studies, and how we applied that knowledge in evolving our designs.

These interfaces entail a conflict between the need for privacy mechanisms and the need for simplicity of representation. We discuss how the conflict arises and some of the difficulties in resolving it. We present the novel tangible interface in detail. Finally, we offer a set of guidelines for the design of audio space systems, and identify areas that require further investigation.

DESIGN CONSIDERATIONS FOR AUDIO SPACES
Empirical data support the hypothesis that audio alone is sufficient to create a usable media space system. Audio has been found to have a primary role in communication [15,17].

* Author's current address: AudioVelocity, Inc., 312 Stearns Hill Road, Waltham, MA 02451, +1.781.894.1402, lisa@audiovelocity.com.
† Author's current address: WhoWhere, Inc., 1675 N. Shoreline Blvd., Mountain View, CA 94043, +1.650.938.4400, sean@whowhere.com.

Somewire's primary characteristics are that it is *persistent, lightweight,* and has *high-quality, spatialized audio.* Persistency refers to communication being available continuously, in contrast to telephone calls which are explicitly started and stopped. Lightweightness refers to the lack of effort required to initiate or end communication; again, the telephone is not lightweight because of the need to pick up the handset, dial a number, and so on.

In support of the importance of the first two characteristics, several office-share studies (e.g., [7]) found continuous open audio to be important to creating and maintaining long-term interaction patterns between colleagues. Gaver [10] pointed out the importance of ambient audio in the workplace for subtly informing people of activities around them. Similarly, Whittaker and colleagues [25] studied informal workplace communication and characterized workplace interactions as one long intermittent conversation, made up of numerous, very short interactions. They predicted that persistent audio and video links could support frequent, brief, lightweight interactions at minimal cost.

The value of high-quality audio has been documented by studies that examined the audio-only condition in media spaces with multiple media. These studies demonstrated the value of providing high-quality full-duplex audio with no transmission lag. For example, Gale [9] found that high-quality audio resulted in faster group task completion times than did low-quality audio combined with video. There is also evidence that low-quality audio adversely affects communication [12].

Spatialized audio uses stereo to create an audio image around the user, like that created by high-fidelity entertainment systems. Spatialized audio is closely related to high-quality audio. Buxton's Hydra system [4], which used small audio-video units for teleconferencing, is particularly interesting in that it allowed users to configure spatially their conference space.

Our use of the term *spatialized* does not imply true three-dimensional (3D) audio, which applies complex signal processing to a mathematical model of the listener's head; such work informed our appreciation of some kind of spatialization, but 3D audio requires headphones to work well. Of interest here is Aoki, Cohen, and Koizumi's [2] audio conferencing system, which has true spatialization, a rich set of audio controls, and a screen interface that supports spatial positioning. That system, however, appears to be primarily a platform for audio spatialization research, rather than a system for workaday use.

One important point is that an audio-only system is considerably less complex, and therefore more practical, than a system that includes video. Somewire is such an audio-only system, and it is described in the next section.

THE SOMEWIRE AUDIO SPACE SYSTEM

Somewire is an audio-based communication system that connects multiple users in separate offices to one another in a manner conceptually similar to a telephone party line or conference call. Connections are continuous rather than transient, and no handset or connection setup is required after an initial configuration.

The goal of the system is to support persistent, lightweight and serendipitous communication among people located in separate physical spaces. Users hear one another's speech and office sounds continuously. This persistency creates and supports serendipitous communication of a kind that rarely occurs with telephone interaction; however, users are free to disconnect themselves from Somewire at any time.

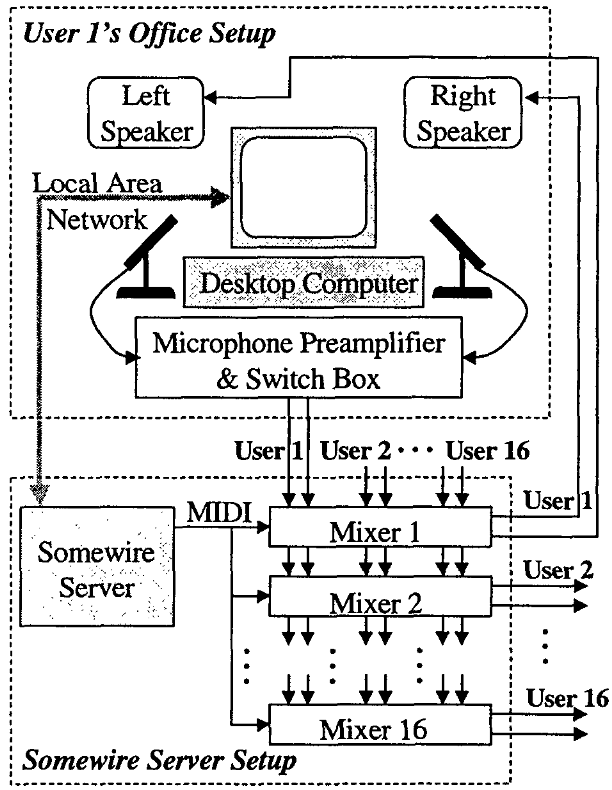


Figure 1. The overall Somewire system for the Faders, Vizwire and ToonTown user interfaces.

Each user's office contains stereo desktop microphones, a pair of speakers, and a hardwired box with an on–off switch that also contains the microphone preamplifiers. The communication process is lightweight. There is no call to place and no switch to throw; users simply speak. Somewire uses analog audio to connect up to 16 users through a central server (Figure 1). The underlying machinery of the system is a set of digital audio mixers, under the control of a networked server [18]. Each mixer is dedicated to creating and maintaining the acoustic space for a single user's office.

The audio is high quality, such that users can easily distinguish one another's voices and clearly hear each other's utterances, even when several users talk simultaneously. The sound quality makes it possible to hear every noise that you might hear if you were sitting in a person's office, including keyboard sounds, telephone conversations, body noises, and background noise.

Unlike other audio conferencing systems, e.g., the one created by Aoki, Cohen, and Koizumi [2], headphones are not necessary in the three Somewire systems that use stereo speakers. With speakers, feedback howl can result if a user's own audio signal is fed back to her system; headphones provide a simple solution to this problem in the fourth Somewire system. The other three versions of Somewire use mixing technology to produce an audio signal that eliminats the local microphones from each office's signal.* Most users work in private offices, so speakers are not problematic for them.

The audio managed by Somewire is not directly connected to the user's computer, although, in two versions, the user controls Somewire via an interface program running as a client application on his computer. The system's model enables widely varying interfaces to operate simultaneously and to interact with one another smoothly. The client programs could modify audio parameters-volume, panning, bass, and treble-for each source sound in the user's local input mix by sending requests to the server. These sound sources are primarily other Somewire users; compact disks and radio broadcasts are also available.

Two of the client application programs, Faders and Vizwire, are screen-based user interfaces. The third, ToonTown, is a physical interface that uses only tangible objects supplemented by auditory indicators. A fourth, Thunderwire, is based on a simpler implementation with just physical controls and headphones.

We shall describe these user interfaces in terms of the system setup, the interfaces' affordances, and users' reactions to the interfaces.

SCREEN-BASED INTERFACES TO SOMEWIRE

Initially, the system designers were predisposed towards conventional graphical user interfaces for controlling Somewire. The first two interfaces we built took this form.

Faders: A Literal Representation

The first interface is **Faders,** which presents a *metaphor of the underlying mixing technology, with sound sources represented by bands of sliders.

In Faders (Figure 2), each user on Somewire is represented by a name and a set of audio controls. The sliders indicate the state of the system in relation to each associated user. *Volume* sets the loudness of each associated remote user's microphone in the local user's mix; *Pun* similarly sets the relative balance of the left and right channels, and *Buss* and *Treble* set these signal characteristics. *The Master Volume* control sets the overall volume of the local user's speakers. The *On-Off* button controls the whole mixer

This literal interface has the advantage of directly presenting controls for the audio parameters that the user can manipulate. It has the disadvantage of requiring that users understand the system's implementation and determine how to manipulate the audio parameters to achieve a desired end.

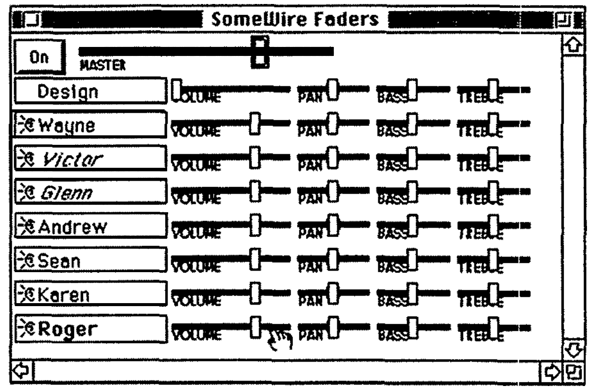


Figure 2. A Faders display. The local user, Roger, is listening to all users except Design', the name denoting a group meeting room.

Vizwire: A Social Representation

The mixer paradigm in Faders provides a *view* into the system machinery itself. **Vizwire,** on the other hand, presents a model that emphasizes the social and physical aspects of audio communication.

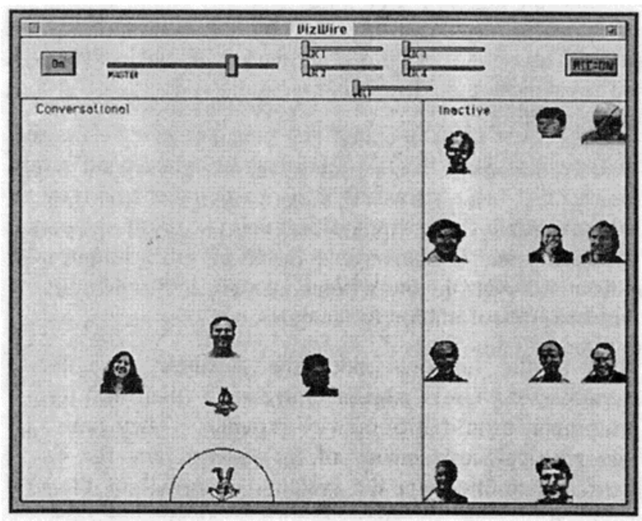


Figure 3. A Vizwire display. The local user, the Rabbit, is talking to four remote users.

In Vizwire (Figure 3), the local user is represented by a fixed icon in a semicircle at the bottom of the left-hand (conversational) region of the display. Each remote Somewire user is represented by a draggable icon on the screen. The Vizwire display is divided into three regions. *The* top *control region* contains overall settings for the system. Remote users whose icons *are* in *the inactive region* on the right side of the display are not included in the local user's mix. Users whose icons *are* in *the conversational region* on the left side of the display are heard in **the** local user's mix. An exception is a remote user whose microphone is turned off. In this case, the remote user **is** not heard and a red slash appears over his icon.a

*At the time, this required racks of sound apparatus. Now the same task could be accomplished much more easily, due to advances in digital audio mixing technology.

The conversational region employs a user-centric presentation. As a remote user's icon moves around the local user's icon, Vizwire moves the remote user's apparent audio position correspondingly. It does this by manipulating the remote user's stereo pan and volume in the local user's mix. Volume is thus represented as a vertical distance between icons, rather than as a linear slider as in the Faders interface. Just as the vertical position of the remote user's icon determines its volume; its horizontal position determines its stereo panning, that is, its left-right relative loudness.

One user can also "whisper" to another—that is, create a temporary private audio space for the two of them—by clicking on the other user's icon.

The advantage of Vizwire's social representation is that it enables users to concentrate on creating a desired social situation. This is a considerable improvement over the Faders interface, which requires users to translate device operations into the corresponding social situation.

User Reactions to Faders and Vizwire

To understand how users perceived the interfaces to Somewire, we conducted interviews after the system had been stable and running continuously for about 6 months. The user community consisted of Somewire project team members and several other researchers, for a total of 17 current and former users.

We formulated two dozen semi-structured questions to elicit users' positive and negative reactions to Somewire; the social influence of use of Somewire; memorable incidents; the way that users operated and conceptualized the system, and typical daily usage. Each user was asked all relevant questions. We also asked users to demonstrate how they operated their Somewire systems and to explain the system's current state. The interviews lasted 10 to 25 minutes; they were videotaped, transcribed, coded independently, and entered into a database for analysis.

The results highlight both the system's strengths and weaknesses. Users almost universally like the form of communication that Somewire provides. They praise both the physical convenience of the system, and the ease of social interaction via the system. One-half of the users, however, are concerned that the system intrudes on their privacy. Not knowing who is listening on the system is a frequent privacy complaint. Users also complain of being occasionally distracted by sounds from their Somewire system, and they find the two GUIs to be awkward to use.

Surprisingly, the hardware on–off switch is the control mechanism that users prefer once the system is configured to their liking. Users do not want to devote screen space to a function that they perceive as being like that of a telephone. An important research idea that emerges from these interviews is the notion of a nongraphical interface—one that is neither screen-based nor computer-based, but rather is like a freestanding physical device.

NON-SCREEN-BASED INTERFACES TO SOMEWIRE

We built two interfaces that were not standard graphical user interfaces: Thunderwire and ToonTown.

Thunderwire: A Physical User Interface

In **Thunderwire**, we eliminated the display and almost all of the user controls.

System Setup

Each user has a pair of desktop microphones, headphones, and a control switch with three settings: off, listen only, and on. There is an on–off indicator light for the microphones, and the sound volume can be adjusted. The overall system design is considerably simpler than the other Somewire systems. Because headphones are used, up to 10 Thunderwire users can be mixed together via a single audio mixer, rather than requiring one mixer per user. Furthermore, there is no software client application, so there is no need for the Somewire server or users' computers to be connected to the Thunderwire system.

Affordances

Thunderwire is like an old-fashioned party line telephone; users share an acoustic space, and they can only control whether their own microphone is active or not. All utterances and sounds from each active user's office or cubicle are heard by all users currently on the system. System use is fluid; people can connect or disconnect from Thunderwire at any time simply by flipping a switch.

Thunderwire is a purely audio medium. Except for the control switch, on–off light and volume control, it has no other visual interface or cues. Connection or disconnection by any user is indicated only by a barely audible click; in fact, there is no way to know exactly who is listening except by asking users to identify themselves.

ToonTown: A Tangible User Interface

Our choice of a tangible user interface for Somewire was greatly influenced by Bishop's marble answering machine [14] and by the emergence of augmented environments as an approach to computationally assisted interaction [24]. Bishop's prototype is an early example of an augmented environment. In the prototype, identification hardware is glued onto ordinary marbles, and a holder that could read a marble's ID is connected to a computer. The marbles could be treated as if they contained voice messages. In Bishop's scripted demonstration, the altered marbles readily afforded message replay, segregation by recipient, and reuse.

ToonTown makes use of similar *active objects*—that is, physical instantiations of computer-based objects that can be manipulated in the same manner as other objects in the physical world. In ToonTown, a user can manipulate physical representations of other users to control their acoustic space. The controls are moved off the screen and into the user's physical space. ToonTown can be described as an *body-syntonic interface*, that is, an interface that allows users to draw on their body knowledge directly [13].

Left pan		S	O	F	T	E	S	T		Right pan	
Left pan					Center Pan					Right pan	
Left pan					Center Pan					Right pan	Info Zone
Left pan		L	O	U	D	E	S	T		Right pan	Assign Zone

Figure 4. ToonTown prototype (top) and the meaning of each position on the board (bottom)

System Setup

The system setup is identical to that used by the screen-based interfaces, except that the display and GUI are replaced by the ToonTown active objects board and pieces. The board is connected to the user's computer and communicates with the ToonTown client software, which in turn controls the Somewire server.

The ToonTown object board and pieces are shown in the photo in Figure 4. This same board is also used in the Logjam system [5]. The prototype object board uses a microprocessor to sense the presence and location of the active object pieces. Each piece is a wooden block, about 1 inch square that contains a unique identification chip. The board has 48 locations arranged in four, rows and 12 columns. Pieces are designed to sit on the board like blocks in a Scrabble™ tray, making contact in two places-the bottom and back of each row.

Tangible representations for the pieces are made from toy cartoon and action figures. The figures can be attached interchangeably to the active object pieces. A writable surface can be affixed to the front of each piece, so users can write labels or reminder notes on the pieces.

Affordances

ToonTown differs from the graphical interfaces in that the user is not represented at all. It feels natural to users to project themselves into the screen when using Vizwire; that is, they are represented by their own icon on the screen and they move other users' icons in relation to their own iconic representation. However, it quickly becomes evident that self-representation is no longer appropriate once the interface moves off the screen and into a physical form; it is not natural for users to have a physical self-representation in addition to their actual physical selves.

The ToonTown object board is shown in Figure 3. Each column on the board, except the rightmost one, represents a spatial location from leftmost to rightmost in pan. Each row represents a volume; the front row is loudest and the back row is softest. The rightmost column is a control area. Placing a piece in the AssignZone causes a list of users to be displayed on the computer screen so that the piece can be assigned to a specific user. Placing a piece in the InfoZone causes an audio segment of the user speaking his or her name to be played, followed by a status message and an optional personalized message. The message is played only when a piece *is* initially placed in this location; the user can interrupt it by picking up the piece.

Audio feedback, in the form of a rising or falling pitch, indicates whenever an object is moved on or off the board.

User Reactions to Thunderwire and ToonTown

The Thunderwire system is robust and has been in use periodically for over a year. Nonetheless, as we report elsewhere [1], the field study clearly suggests user interface improvements. Users would prefer to know who is present in the audio space, and to have an automatic mechanism for turning off microphone input during an incoming telephone call. Furthermore, they would like the ability to set up two-way, private conversations.

With respect to ToonTown, users consider its tangible interface to be a highly engaging means of interaction. When they first see the object board and characters, several users have remarked that they feel irresistibly drawn to play with the pieces, and users enjoy being represented by a character. Because only a single prototype board exists for interacting with the Somewire system, we have not rigorously investigated usability and collaborative use.

PRIVACY, REPRESENTATION, AND PRESENCE

Previous media space research (see Bly, Harrison and Irwin [3] for a contemporaneous summary) has highlighted the importance of privacy. Privacy is of concern to us, in the design of the Somewire system and in the nature of users' social experiences with one another via Somewire.

In particular, the persistency and lightweightness that enables Somewire's key benefit, casual communication, also makes privacy violations almost inevitable. Privacy violations, group norms, and related social effects are addressed at length in a previously published study [1]. Here, we look at Somewire's privacy model, the implications for user interactions with the system, and the need for interface mechanisms that can provide users with an awareness of who else might be using the audio space.

Privacy Model

Somewire's privacy model emphasizes each user's control over his own acoustic space, and over the information about him that is available to other users. A user can control only what he hears and who can hear him; he cannot control the volume at which other people hear him, or his spatial position in another user's acoustic space.

This approach ensures that the system does not allow any user to be made audible to any other user without express action by both of them. This constraint is valuable in fostering trust in the system machinery. However, it turns out to complicate the representation of system configuration, and therefore to complicate a user's ability to understand and control that configuration.

Individual Control and Lack of Symmetry

All the Somewire interfaces except Thunderwire use a personal point of view—each user on the system controls his or her own acoustic space, moving people to the left or right, making each person louder or softer. The representations ware asymmetric in that user A might place user B to her left, and user B might place user A on her left as well—there is no global coordinate system.

A single global display would have eliminated the asymmetry, at the unacceptable cost of also eliminating each user's control over her own space. Furthermore, the privacy model does not support sharing of configuration information about users other than the two involved in any specific active or inactive state of connection. Even if this were not the case, it is not clear how the multi-way configurations among 16 users could be clearly presented.

In Thunderwire, point of view is irrelevant because there is no representation or control over anything other than the status of a user's own microphone. This solution works remarkably well.

Presence Awareness

Users of all the different interfaces have expressed a strong desire to know who is present on the system and thus potentially listening to them. Knowing whom they had *allowed* to listen is insufficient to forestall perceived violations of personal privacy.

Somewire can only indicate the status of a user's microphone. For example, in Vizwire a red slash indicates that a user's microphone is muted. In ToonTown, a user can determine whether another user's microphone is open by placing that user's piece in the InfoZone space on the board. Somewire cannot determine or indicate a user's actual attentiveness to Somewire, although it is technically feasible to indicate who has recently spoken aloud and who is physically present in their offices.

TANGIBLE USER INTERFACES FOR COMMUNICATION

ToonTown focuses on using tangible interfaces for communication, and users are represented as tangible objects. This section discusses social implications of representing people as objects, as well as how ToonTown is related to recent work in tangible user interfaces.

Social Implications of Representations of People

The ToonTown interface explores a novel use of active objects, where the objects represent people rather than software tools or computer-stored media.

Such a representation of people has social implications. For the ToonTown interface, each user selects his representations for the other users. Thus, it is possible that a person would not like the representations that other users select for her. In Vizwire, by contrast, each person is represented throughout the system with a self-selected icon. In one multi-user chat system studied by Schiano and White, women were more conscious of their representation than were men, and they were more concerned about having control over their representation [16].

Other social implications can result from using physical objects to represent people. How will users feel about picking up people by their heads? Would they 'flick' people off the board when they don't want to listen to them? What kinds of interactions would become acceptable practice?

Related Work

Our exploration of tangible interfaces for communication is a significant contribution to research in tangibility. Fitzmaurice, Ishii and Buxton [8] laid out a taxonomy for tangible interfaces, based on their graspable bricks. This taxonomy identified numerous aspects of tangible interactions, but concentrated on the object's interaction with the technology and so cannot be generalized to communication between users or to a specific application such as control of a communication system.

Ishii and Ullmer [11] reported on their metaDESK system, in which a flat display surface contained hardware for optically recognizing and tracking the location of physical objects on the display surface. For instance, metaDESK users could interact with a displayed campus map by moving an object shaped like a specific building. MetaDESK's use of semantically meaningful objects is similar to ToonTown's use of cartoon figures, but metaDESK did not involve communication, representations of people, or a virtual space shared by multiple users.

MediaBlocks [22], the successor to metaDESK, is similar in appearance and affordances to the ToonTown object board and pieces. However, the blocks referred to digital media content rather than to people, and the blocks did not indicate their contents. The mediaBlocks system extends beyond its ToonTown-like grid board to a range of media containers and operations, such as sequencing video clips. However, it did not address multi-user communication or the representation of people in tangible interfaces.

MediaBlocks suggests a possible extension to ToonTown, in which ToonTown users could add content to pieces that are not attached to any given user. For example, audio reminders could be represented by small alarm clocks.

CONCLUSIONS

We drew two sets of conclusions from our experience with, and study of, these four interfaces. The first set are design guidelines for audio space systems. The second set are questions for future research on audio space systems.

Design Guidelines

- **GUI interfaces are a poor choice for audio spaces.**
  In retrospect, it is not surprising that a graphical user interface is not optimal for interacting with an auditory experience. Audio communication does not demand visual attention. Furthermore, an audio space works like a utility and thus calls for a simple interface. With Thunderwire, we took simplicity too far by eliminating all forms of control and display. Some kind of tangible representation, building on a simpler version of the ToonTown model, might be the more appropriately balanced interaction mechanism.
- **Users do not need control over audio localization, or, by inference, over other audio attributes.**
  We were inspired by user reactions to build successive interfaces with less and less control for audio characteristics. Bass and treble controls available in Faders were dropped from subsequent interfaces; left–right panning, which is a prominent aspect of Vizwire and ToonTown, was dropped with Thunderwire.

  Our experience with localization is particularly instructive. With each of the interfaces that represented position—left–right relative position in Faders and Vizwire, and grid location in ToonTown—localization is a source of user confusion. Furthermore, when it was removed altogether in Thunderwire, users did not complain about its absence.

  This elimination of functionality leaves a system distinct from audio conferencing systems, which typically feature numerous audio controls [2]. Our discovery that here too, less is more, is particularly salient now that digital signal processing is becoming commonplace on personal computers. Designers of interfaces to general-purpose audio environments should be encouraged to resist the temptation to make every possible control available to the user.
- **Awareness of other users' presence is desirable.**
  One way to represent presence in the audio itself is through the use of auditory feedback. Cohen's 'Out to Lunch' [6] system provided auditory feedback about the presence of other workgroup members. Users could get abstract information about other users' activities (e.g., audio feedback of keyboard and mouse activity), but could not converse. The low-disturbance audio explored by Smith and Hudson [19], where users can hear who is speaking without hearing the words themselves, is another way to indicate whether other users are active in an audio space.

Matters for Further Investigation

- **What is audio quality's role in audio spaces?**
  Much of Somewire's attractiveness and utility resulted from its clean audio signal. Based on our experiences, we believe that the use of stereo microphones and speakers affords a spatialization effect that is critical to creating the illusion of an acoustic space. The space illusion is completely distinct from specific control over localization, which is not a necessary feature.

  It may also be that the role of audio quality in a successful audio space is less significant than we originally thought; Strub's study of two-way radio use over a weekend by groups of teenagers [21] showed that even low-quality, persistent audio could enable behavior indicative of social closeness.
- **How can audio spaces scale beyond a workgroup?**
  The Vizwire interface provided a single acoustic space and contained one representation of each user. However, what if users want to be in multiple spaces at once (e.g., to listen to multiple conversations)? One approach would be to have doppelgangers, that is, multiple representations of users.

  The voice-loop systems used by space mission controllers provide additional insight into how we might meet this need. In these systems, multi-layered, complex audio spaces are made usable through constraints on who is allowed to speak to whom, on the use of foreground and background volume levels, and on language use. The applicability of these approaches when the audio system is not the user's primary work task is unknown [23].

Through our work in creating a variety of interfaces to the Somewire audio space system, we have explored representation and control in collaborative shared audio environments. We have not made an exhaustive inquiry, however. We look forward to seeing how audio space interfaces will evolve, and how they will converge with the widespread availability of Internet-based audio communication.

ACKNOWLEDGMENTS

The authors thank the rest of the Somewire team and users, especially Glenn Edens for architecting Somewire and Roger Meike for creating the client and server software infrastructure. Wayne Burdick and Don Charnley built the Somewire hardware and the Vizwire software, respectively.

Bill Verplank crafted the board and pieces, which were instrumented by Scott Wallters and others. Jonathan Cohen provided helpful input to the design of ToonTown. We also thank the many reviewers of earlier versions of this paper, and Lyn Dupré for her editing assistance.

REFERENCES

1. Ackerman, M., Hindus, D., Mainwaring, S. and Starr, B. Hanging on the 'wire: A field study of an audio-only media space. *ACM Transactions on Computer-Human Interaction,* March 1997, 4(1):39-66.

2. Aoki, S., Cohen, M. and Koizumi, N. Design and control of shared conferencing environments for audio telecommunication using individually measured HRTFs. *Presence*, 1994, 3(1):60-72.

3. Bly, S.A., Harrison, S. R. and Irwin, S. Media spaces: Bringing people together in a video, audio, and computing environment. *Communications of the ACM*, January 1993, 36(1):28-47.

4. Buxton, W. Telepresence: Integrating shared task and person spaces. In *Proceedings of Graphics Interface'92,* (Vancouver Canada, May 1992), Morgan Kaufman Publishers, San Francisco, 123-129.

5. Cohen, J., Withgott, M. and Piernot, P. Logjam: A tangible multi-person interface for video annotation. To appear in *Proceedings of CHI'99* (Pittsburgh PA, May 1999), ACM Press.

6. Cohen, J. Out to lunch: Further adventures monitoring background activities. In *Proceedings of ICAD'94, International Conference on Auditory Display.* (Sante Fe NM, November 1994), Sante Fe Institute, 15-20.

7. Dourish, P., Adler, A., Bellotti, V. and Henderson, A. Your place or mine? Learning from long-term use of video communications. *Computer Supported Cooperative Work,* 1996, 5(1):33-62.

8. Fitzmaurice, G., Ishii, H. and Buxton, W. Bricks: Laying the foundations for graspable user interfaces. In *Proceedings of CHI'95* (Denver CO, May 1995), ACM Press, 442-449.

9. Gale, S. Human aspects of interactive multimedia communication. *Interacting with Computers*, 1990, 2(2):175-189.

10. Gaver, W. W. Sound support for collaboration. In *Proceedings of European CSCW'91* (Amsterdam The Netherlands, September 1991), Kluwer Academic Publishers, Boston, 293-308.

11. Ishii, H. and Ullmer, B. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of CHI '97* (Atlanta GA, March 1997), ACM Press, 234-241.

12. O'Conaill, B., Whittaker, S. and Wilbur, S. Conversations over video conferences: An evaluation of the spoken aspects of video-mediated communication. *Human-Computer Interaction*, 1993, 8(4):389-428.

13. Papert, S. *Mindstorms: Children, Computers and Powerful Ideas*. Basic Books, 1980.

14. Poynor, R. The hand that rocks the cradle. *I.D. Magazine*, May/June 1995, 60-65.

15. Rutter, D. R. *Communicating by Telephone*. Pergamon Press, New York NY, 1987.

16. Schiano, D. and White, S. The first noble truth of cyberspace: People are people (even when they MOO). In *Proceedings of CHI'98* (Los Angeles CA, April 1998), ACM Press, 352-359.

17. Sellen, A. J. Remote conversations: The effects of mediating talk with technology. *Human-Computer Interaction*, 1995, 10(4):401-444.

18. Singer, A., *et al.*, Methods and systems for creating a spatial auditory environment in an audio conference system. Patent pending, United States Patent Office.

19. Smith, I. and Hudson, S. E. Low disturbance audio for awareness and privacy in media space applications. In *Proceedings of Multimedia'95* (San Francisco CA, November 1995), ACM Press, 91-97.

20. Stifelman, L. Augmenting real-world objects: A paper-based audio notebook. In *CHI'96 Conference Companion* (Vancouver Canada, April 1996), ACM Press, 199-200.

21. Strub, H. ConcertTalk: A weekend with a portable audio space. In *Proceedings of INTERACT'97* (Sydney Australia, July 1997), Kluwer Academic Publishers, Boston, 381-388.

22. Ullmer, B., Ishii, H. and Glas, B. mediaBlocks: Physical containers, transports, and controls for online media. In *Proceedings of SIGGRAPH'98* (Orlando FL, July 1998), ACM Press, 379-386.

23. Watts, J., Woods, D. D., Corban, J. M., Patterson, E. S., Kerr, R. L. and Hicks, L. C. Voice loops as cooperative aids in space shuttle mission control. In *Proceedings of CSCW'96* (Boston MA, November 1996), ACM Press, 48-55.

24. Wellner, P., Mackay, M. and Gold, R. (eds). Computer augmented environments. *Communications of the ACM*, July 1993, 46(7) (entire issue).

25. Whittaker, S., Frohlich, D. and Daly-Jones, O. Informal workplace communication: What is it like and how might we support it? In *Proceedings of CHI'93* (Amsterdam The Netherlands, April 1993), ACM Press, 131-137.

Tangible Progress: Less is More in Somewire Audio Spaces (Page 104)

Andrew Singer, Debby Hindus, Lisa Stifelman, and Sean White

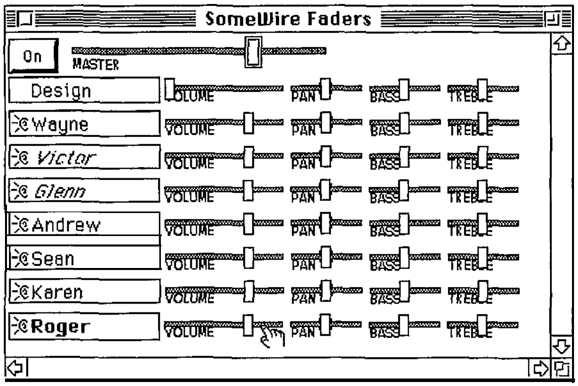


Figure 2. A Faders display. The local user, Roger, is listening to all users except Design', the name denoting a group meeting room

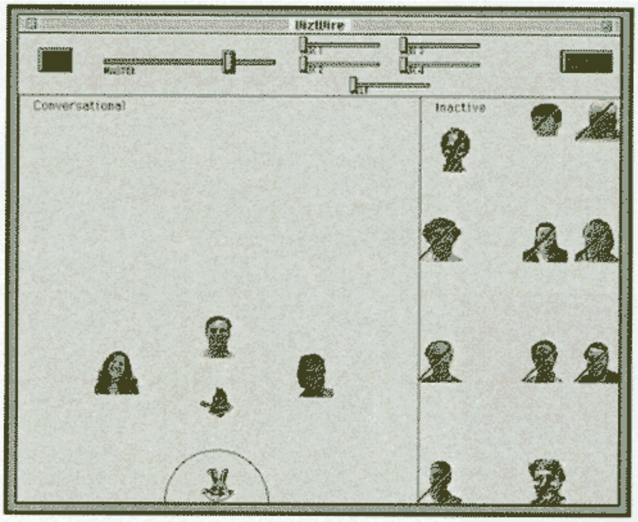
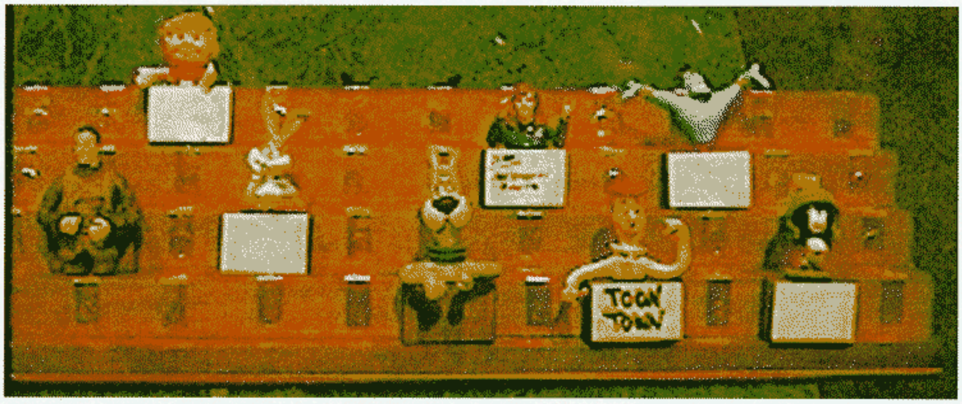


Figure 3. A Vizwire display. The local user, the Rabbit, is talking to four remote users



Left pan		S	O	F	T	E	S	T		Right pan	
Left pan					Center Pan					Right pan	
Left pan					Center Pan					Right pan	Info Zone
Left pan		L	O	U	D	E	S	T		Right pan	Assign Zone

Figure 4. ToonTown prototype (top) and the meaning of each position on the board (bottom)