Talking In Circles: Designing A Spatially-Grounded Audioconferencing Environment

Roy Rodenstein, Judith S. Donath
Sociable Media Group
MIT Media Lab
20 Ames Street
Cambridge, MA 02142
{royrod, judith}@media.mit.edu
http://www.media.mit.edu/~royrod/projects/TIC

ABSTRACT
This paper presents *Talking in Circles*, a multimodal audioconferencing environment whose novel design emphasizes spatial grounding with the aim of supporting naturalistic group interaction behaviors. Participants communicate primarily by speech and are represented as colored circles in a two-dimensional space. Behaviors such as subgroup conversations and social navigation are supported through circle mobility as mediated by the environment and the crowd and distance-based attenuation of the audio. The circles serve as platforms for the display of identity, presence and activity: graphics are synchronized to participants' speech to aid in speech-source identification and participants can sketch in their circle, allowing a pictorial and gestural channel to complement the audio. We note user experiences through informal studies as well as design challenges we have faced in the creation of a rich environment for computer-mediated communication.

Keywords: Computer-mediated communication, audio, speech, drawing, representation, media space, interaction design, multimodal interfaces, multicast, social navigation, gesture

INTRODUCTION
Communication is one of the primary applications of computing. Electronic mail has become ubiquitous in certain sectors, with synchronous computer-mediated-communication surging in the last decade through networking improvements and critical mass of the online population. Chat, whether purely textual or accompanied by graphics, has ridden the growth of the World-Wide Web to become a popular medium for social interaction. Traditional chat environments, however, are limited by the physical and expressive bounds of typing as input for synchronous communication. Disparities in participants' typing ability and expertise with particular writing conventions of the online medium, as well as the lack of traditional emotional cues such as tone of voice, are among these limitations.

Audio-based communication ameliorates some of these issues, allowing voice interaction to leverage users' experience with spoken conversation. Audio spaces have demonstrated clear potential for fostering rich social interactions [1]. However, though speech is a very natural form of communication and allows a great range of expression, audio-only spaces place multiple conversations into a single audio stream, serializing speaker interactions, and establish user presence only during their transitory speech [26]. We address these issues through the use of simple graphics that display presence, enable parallel conversations, and reveal the evolving interactions within the social space.

With the aim of supporting rich, naturalistic social interaction we have developed *Talking in Circles,* a computer-mediated social environment in which speech is the primary communication channel and graphics convey important expressive and proxemic information. Our environment is a two-dimensional space within which participants are represented as circles of various colors, as shown in Figure 1. Participants can draw on their circle, and the system graphically shows when each is speaking. In addition, users hear those they move close to more clearly than others farther away. This property has supported naturalistic behaviors such as users approaching those they are interested in conversing with, forming conversational subgroups of several users located near each other, and mingling by moving around within the space, listening for topics of interest and moving to join particular groups.

Participants' circles thus serve as indicators of presence and as cues about membership and activity in various conversations, one way our system addresses the problem of observing the state of the media space. The circles are also used as platforms for graphical display. The system visually renders each participant's speech activity on their circle, making it easy to identify individual speakers and differentiate among several overlapping speeches. It also allows users to draw on their own circles, complementing the audio channel with pictorial communication. The screen-shot in Figure 1 shows these graphical features at one instant in time in a *Talking in Circles* chat session.

In the following section we review work which relates to our design for *Talking in Circles*. We then discuss the process of our system's interaction design and analyze it in the context of relevant studies. Next we briefly describe the implementation, noting our responses to technical challenges we encountered. We conclude by noting our ongoing work in this area as well as future directions.

RELATED WORK

Many systems have been used to research computer mediation of audible communication between people. In general these are geared toward computer-supported collaborative work or focus on particular modalities. *Thunderwire* is one audio-only media space studied over several months [10]. Although the system did not foster much work-related communication, it was very successful as a sociable medium based on such criteria as informality and spontaneity of interactions. Interesting usage norms evolved as a result of the lack of visual feedback to deal with finding out who was listening on the system and to indicate lack of desire to converse at particular times. Our work leverages the pliable, sociable quality of the audio medium and uses a graphical interface to resolve participants' lack of knowledge about the space's membership and member interests.

Sun's *Kansas* system uses videoconferencing within small windows and a distance threshold for audio [13]. Geared toward distance learning, it employs a screen-sharing approach to complement videoconferencing with various applications. *FreeWalk* also used video, mapping it onto a flat surface and placing participants in a three-dimensional environment [19]. *FreeWalk* succeeded in fostering certain social behaviors, such as following a participant from afar. Audio was faded based on distance, though not based on the direction participants faced, leading to difficulty in speaker identification in some cases. The excellent infrastructure provided by the more general *MASSIVE* project had similar problems in some of the interfaces developed for it, due to limited representational feedback for participants' speech [8]. *Talking in Circles*' use of animation synchronized with users' speech resolves these sound-source-identification problems.

In terms of point-of-view, both *MASSIVE* and *FreeWalk* focused on close-up views of figurative user representations, differing from our work's emphasis on highly abstract representation and visual overview of the space. Though *MASSIVE* offered flexible viewpoints there were sometimes resultant mismatches in behavior, such as users with disparate views walking right "through" others' bodies. We chose as a tradeoff a single cohesive viewpoint, and address the issue of personal physical boundaries in the section on circle motion.

In the realm of graphical text-based chat, systems such as *The Palace* [21] and *Comic Chat* [16] use typed communication enhanced with changeable displays of avatars for participants. *Babble* shows participants as colored dots which converge toward the center of a circle as they actively take part in a specific chat, selected from a list of topics [7]. Previous work in our group includes *ChatCircles*, a text-based two-dimensional chat space which explored abstract representation of participants as circles as an alternative to the broadly caricatured feel of traditional avatars [32]. *ChatCircles* displays participants' typed messages on their circle, whose expansion thus signals activity. The text is visible only to participants within a threshold distance from the sender.

Our work on *Talking in Circles* makes several contributions to research in computer-mediated communication, including a novel approach to resolving collisions between participants' graphical representations, graphical aids to speech-source indentification, and prominent access to lightweight multimodal communication. More broadly, our work is characterized by a willingness to explore non-traditional representations, exploit rich yet low-overhead interface affordances, and focus specifically on *social*, as opposed to task-oriented, communication. We now discuss this work in detail.

SOCIAL INTERACTION DESIGN

A first issue we dealt with in creating an audio-based communication space was helping users map the voices they hear to the circles representing the respective participants. That is, we wanted to move the user experience from disembodied voices to a more cohesive perception of fellow participants. We thus use a bright inner circle displayed inside their darker circle to represent the instantaneous energy of the speech from a particular participant. Figure 1 shows that Sandy and Al are speaking, with their inner circles' size showing how loud their speech was at that instant. Thus, natural pauses in conversation, which leave only particular circles showing speech activity, make this cognitive matching problem much simpler. Distance-based spatialization, discussed below, also helps, as speech from circles farther from the user's circle sounds fainter. Finally, identity cues such as learning a participant's voice or their circle's labeled name also resolve matchings.

We conducted an informal test of the graphical feedback provided by the dynamically-changing bright inner circle. Six subjects were shown two circles with non-identifying names, equidistant from their own circle and at equal audio volume. The two test circles each played a different RealAudio news stream, and we asked participants about their experience in trying to match the two speakers they heard to the correct circles.

Although this scenario is challenging, with constantly overlapping speech and no individuating cues for the circles, all subjects successfully matched each stream to its corresponding circle within a few seconds. Though simultaneous speech was at first confusing, the subjects mentioned that the occurrence of short natural pauses in the speech soon made the matching apparent, as only one voice was heard and only one circle was bright. A similar situation happened when one speaker said something loud causing one bright inner circle to grow visibly larger than the other. In general, the subjects said the speech-synchronized updating of the bright inner circle's helped them differentiate and identify the speakers by highlighting the matching rhythms of the speech.

We also experimented with graphical display of recency of participant activity, allowing the bright inner circle to fade out slowly over time when a speaker stopped speaking. This provided a slightly enhanced short-term history but it interfered with the real-time feel of the inner circle's rhythm. In addition, we noted that recency of activity is not necessarily equivalent to availability. Although we think availability display is useful, accurate automated detection of participant availability is infeasible while manual control of availability status is also unreliable. As Ackerman found local disruptions can cause frequent changes in listening or speaking availability of users without their remembering to turn off their microphone, even when it resulted in unwanted eavesdropping [1]. Similarly, lack of speech by a user for several minutes does not guarantee that they are not still listening. The system therefore does not currently attempt to display availability status.

Spatialization

While the system supports various user capabilities, some are supported directly while others arise out of a combination of modalities. The most salient behavior is that of circle motion as an indication of interest and membership in a conversation. As Milgram notes, rings are a naturally-emerging configuration of people engaged around a common activity [18]. In *Talking in Circles*, as in face-to-face situations, standing close to someone permits one to hear them clearly while also reducing distraction from other sources farther away, due to distance-based audio attenuation.

This natural tendency toward physical alignment, besides being a functional conversational feature and serving to a limited extent the role of gaze, has additional benefits. It allows other participants to view the formation of groups or crowds around a particular discussion, letting them gauge trends in participants' interest and advising them of conversations that are potentially interesting. Crowd motion does not necessarily require explicit attention; as in real life large gatherings stand out, and can continue to draw people as users notice the traffic and wonder what the fuss is about.

An additional important benefit of this crowd motion is simply the vitality with which it imbues the space. Whyte remarks on the fact that the biggest draw for people is other people, and notes the popularity of people-watching as a form of triangulation— simply stated, a stimulus source which can be observed by multiple members of the population, potentially giving rise to conversation between strangers. The grounding in a 2D space may also bring in features such as traveling conversations, where conversants move across a space to find a comfortable spot along the perimeter [33]. As in real life, it is possible, with some effort, to be near a particular speaker but attend to another, or to stand between two groups and attend to both conversations.

Selective attention, enabled by the physical grounding and audio attenuation, also provides some of the benefits related to the 'cocktail-party effect' [5]. Though audio from those one is closest to is heard most clearly, nearby conversation can be heard more softly within a certain distance threshold. This helps a user concentrate primarily on the conversation group they have joined while preserving peripheral awareness with the possibility of 'overhearing,' such that the mention of a name or keyword of interest can still be noticed. Thus, social mingling is fostered, as participants can move between subgroup conversations as their interest changes or move to an unoccupied physical space and start their own conversation.

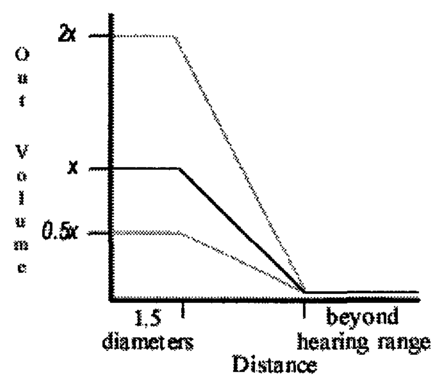

Figure 3: Output volume as a function of distance from a speaker, for input volume *x* (black line), *0.5x* and *2x*. The hearing range is 5 diameters.

In order to allow clear audio for participants in a conversation, no audio fading is done within a distance of 1.5 diameters from the center of each speaker's circle. This allows participants located next to or very close to each other to hear the full volume of speakers' speech, while we still perform fading for circles in conversations farther away. Figure 3 plots the shape of the audio-fading function to show how output volume varies by distance from a speaker. The function remains the same but is parameterized by the instantaneous input volume, as shown by the upper and lower lines in the figure. This modification to the spatial rules of our environment preserves the positive qualities of audio fading but helps members of a conversation hear each other clearly; our focus on spatial grounding is always rooted in fostering a sociable space. Though detailed user control over fading parameters could be beneficial, such as in the case of a very widespread conversation group, customizing the physical rules of the space can lead to inconsistent user experiences [29] as well as unnecessary GUI clutter [28].

The distance threshold for audio to be heard, currently 5 diameters, serves multiple functions. Naturally, it aids performance optimization by obviating the need for audio playback for clients beyond the threshold. The major benefit, however, is letting the user know that they cannot hear someone, as activity by those beyond the hearing range is rendered as a hollow circle. For example, screen shots of a *Talking in Circles* chat from the screen of participant Al, the blue circle, shows he has moved from a conversation with Andy and Helen in Figure 1 to one with Josh and Yef in Figure 2. The hollow orange inner circle shows that Al is now beyond the hearing range of Sandy. Since the hollow circle still indicates speech, however, a participant can still note a flurry of activity even if they cannot hear it, and can move closer to see what the discussion is about if they so desire. The audio threshold is symmetric, such that if user X is too far to hear user Y, Y is also too far to hear X. This feature lets a user easily find a spot where they cannot be heard by a certain group, by noting when their inner circles appear hollow. Thus, as in a real cocktail party, one can move to the side to have a semi-private conversation, although this privacy relies only on social rules and is not enforced or guaranteed by the system. These interaction possibilities address some shortcomings of video-mediated conversation, such as the lack of a "negotiated mutual distance" and of a sense of others' perception of one's voice [27].

One group of visitors to our laboratory who tried the system suggested that, in their corporate setting, they would be interested in private breakout rooms for a couple of participants each, as well as a larger full-group meeting room. Although *Talking in Circles* can easily be adapted to support such a mode, our focus on a purely social space makes relying on existing social behaviors more interesting to us than technical enforcement of boundaries. A related concern is that of rudeness or other undesirable behavior by participants. Once again, the system's rich interaction design can support emergent social mores that help sort things out; just as people can move closer to conversations or people they are interested in, they can move away from conversations which become uninteresting or people who show hostility.

Beyond these pragmatic features of distance-based audio attenuation, other potential sociable applications exist. For example, with a low audio distance threshold the popular children's game of "telephone" is playable, in which a large circle is formed by all attendees and a short phrase or story is whispered from person to person around the circle, becoming increasingly distorted, until it gets back to the originator and the starting and ending phrases are revealed to everyone. Though not always desirable, yelling across the room to get someone's attention or to say something to everyone present is also possible.

Circle Motion

Overlap of participants' representations is a problem some graphical chat systems, such as The Palace, have faced. We have observed that avatars partially obscuring each other, whether intentionally or accidentally, can arouse strong responses from those caught underneath. Visibility, as well as bodily integrity of one's representation, are important social factors in graphical chats. *Talking in Circles* deals with both of these factors by not permitting overlap, simultaneously preserving visibility of participants' presence and of their drawing space.

A circle's motion is stopped by the system before it enters another circle's space. In this situation, a participant can drag the pointer inside the circle blocking their path and their own circle follows around at the outer edge of the circle that is in the way, which provides the feel of highly responsive orbiting. Thus, at close quarters, participants still preserve their personal space and can move around in a manner which provides a certain physical interaction with other participants, an attempt at enhancing the feeling of being in a crowd the system provides. Swift motion across large areas is still immediate, as obstructing the participant with all circles along the way to the new location would make for a cumbersome interface. As always, our aim is to leverage spatial grounding with a primary focus on social interaction design, hence our differing policy for motion at close quarters versus over greater distances.

Drawing and Gesture

One major benefit of audioconferencing, of course, is that it frees the hands from being tied to a keyboard. This freedom can be employed to run *Talking in Circles* on a keyboardless LCD-display tablet, with the added benefit of using a more pliant pen for input instead of a mouse, or on a wearable computer. Unlike traditional chat systems we need not display large amounts of text, which takes up a lot of screen real estate, resulting in great freedom in maximizing the potential usage of the space and the graphical area marked off by each circle.

Since the circles' interior space is used only momentarily during speech, this space can be used for drawing. Though the space on one's circle is limited, it is large enough for diagrams, bits of handwriting, and so on. Drawing strokes appear in bright white, visible even over the graphical feedback during speech, and fade away in 30 seconds. Although this makes long-lasting sketching more difficult, a tradeoff worth noting, our design intent is akin to letting people at a cocktail party use a napkin to share sketches on, and obviates potential distraction from cumbersome drawing controls. The relatively fast refresh rate keeps the drawing space available, which is important for drawing to be useful for gesturing.

Drawing faces is a natural tendency, and it is particularly inviting given the circular shape of the user's representation. The circle's space is enough to permit much more than simple emoticons, and even drawing-unskilled users immediately took to writing short phrases and drawing faces. Combined with moving one's circle, drawing a face can be targeted at a particular user both by drawing the face as facing in that user's direction and by moving toward that user with the face showing on one's circle. Coordinating motion with drawing has been popular with users, such as drawing a face with the tongue sticking out and moving quickly up and down next to the intended viewer, enhancing the facial expression with bodily motion.

Shared drawing is also useful for showing explanatory diagrams [30], which Isaacs and Tang note as a user-requested capability in their study [11], for certain kinds of pointing, and potentially for other meta-conversational behaviors such as back channels. These uses are important in creating a social space since studies of telephone conversations have found reduced spontaneity and increased social distance compared to face-to-face discussions [24]. Employing drawing, confusion can be indicated not just by explicit voicing but by a question mark or other self-styled expression on one's circle.

As the system is used in various environments we are very interested in studying the development of novel drawing conventions and gestures for conveying various data. We have already observed novice users effect floor control, for example, by displaying an exclamation mark in their circle upon hearing something surprising, or simply by shaking their circle a bit to indicate they have something to say. Again, although voice by itself is useful in these tasks to some extent, both audio-only and videoconferencing studies have found complex tasks such as floor control to be less effective than can be done in face-to-face communication. Thus, the complementary combination of voice, circle motion and drawing is aimed at overcoming some of these limitations.

In order to make the pictorial modality more accessible, we also provide a set of clickable icons that display drawings in the user's circle, similar to the availability of preset graphics in The Palace. As shown in figure 2, the system currently includes a question mark and exclamation point, as well as expressions indicating happiness, humor, surprise and sadness. The drawings available on the icon bar are standard graphics files editable in any graphics editor, and the drawings the system includes can be removed or modified or new ones added simply by putting them in the *Talking in Circles* directory. The ready access to showing these iconic drawings on one's circle and the ability to customize this set of drawings makes the pictorial channel more available than requiring the user to draw everything from scratch each time. The icon-bar drawings can be clicked on in sequence and are updated immediately, which allows for higher-level expressive sequences such as pictorially sticking one's tongue out while making a humorous remark, followed by displaying the winking face and then the smiling face.

Lastly, drawing can of course be used strictly for doodling, whether out of boredom or to accompany music one is listening to, and for other purely aesthetic ends. Individuals' use of their drawing space —whether they draw constantly or rarely, make abstract doodles, draw faces or words— may provide others a sense of the person's identity. Besides being an expressive channel, these behaviors serve as another form of triangulation, giving participants a spectacle to watch and gather around..

SYSTEM IMPLEMENTATION

The requirements for *Talking in Circles* focused on full-duplex audioconferencing between a substantial number of simultaneous users, where we defined substantial as between a dozen and 20. Even experimentally, we were interested in low-latency audio, as lag is known to be detrimental to the use of speech for social interaction, for example leading to greater formality [20]. However, we were also interested in creating a system that could have as wide a user base as possible, important given our focus on social applications as well as to facilitate wider, extended study of the system's use. This meant that we could not use proprietary broadband networks or high-speed LAN's, as previous systems have typically done.

We initially looked at designing for the internet as a whole, but this proved intractable, as even highly compressed protocols such as RealAudio occasionally suffer from unpredictable network delays and must pause to rebuffer [23]. Next we implemented the system using the Java Sound API [12], but measured end-to-end lag at two to three seconds for machines on the same high-speed LAN subnet. Finally, we settled on adapting RAT, the Robust Audio Tool, an open-source audioconferencing tool from University College, London [31]. RAT uses the MBONE, the internet's multicast backbone, for network transport [25]. Thus we avoided inefficient strictly client-server and peer-to-peer architectures. Each client sends its audio only once and it is then multicast to the other clients.

We modified RAT to support participant state (x/y location, circle color, instantaneous audio energy, and so on) and measured end-to-end lag at approximately 0.3 seconds, considerable but not detrimental unless participants can also hear each other directly [14]. In addition, the bottleneck in our current implementation is screen redraw, as discussed below, and accordingly we have noticed no substantial performance degradation when varying the number of users from one to eleven.

Although the audio code, including compression/ decompression and MBONE transport, is written in C, we maintained the user interaction portion in Java using the Java Native Interface in Sun's Java 2 platform. For example, computations including instantaneous audio energy, background noise suppression and logarithmic normalization to map the energy value onto the circle's area are performed in C, and the bright inner circle (see figures 1 and 2) is then updated several times per second in the Java component. Lag was a problem with Java's mouse-motion reporting during freehand drawing, which we adequately resolved using Bresenham's line-interpolation algorithm. Audio bandwidth use is also moderate at 5KB/s per client.

To summarize the data used by our system, each participant multicasts the following:

- circle x/y coordinates
- freehand drawing
- icon selection
- instantaneous audio energy
- speech/audio

The interface is rendered from these features, and all participants' displays share:

- circles' location
- circles' drawing/icon display
- participants' audio

Finally, the local user's relative location produces a subjective rendering of:

- speech volume (audio)
- speech rendering (graphics)

CONCLUSIONS AND FUTURE DIRECTIONS

Approximately thirty people have used the system for periods up to half an hour, and we are working to deploy it for broader testing by a group of steady users. Reactions have been extremely encouraging, with users reporting enjoyment of the cocktail-party-like environment, the drawing capability, and the graphical feedback for speech. The system's fluid representation of participants' speaking and moving about to converse with different groups was cited as the most satisfying aspect, though a shortcoming users noted is slower update rates when multiple people draw or speak while moving around a lot.

In designing *Talking in Circles*, we have strived for a rich medium for communication along dimensions including interactivity, that is, responsiveness, and expressiveness, or "multiplicity of cues, language variety, and the ability to infuse personal feelings and emotions into the communication" [3]. While speech has been found to have reduced cognitive load compared to text generation [15] and to be the key medium for collaborative problem-solving [4], as Chalfonte found text still has certain advantages. For structured data, such as URL's, text is clearly superior to speech due to its permanence and precision. Further, text-based CMC supports threading more so than face-to-face communication [17], which can add fluidity and variety to conversation. Though the ability to draw and do limited handwriting in *Talking in Circles* can help in some cases, textual communication can nevertheless add to the set of useful channels at users' disposal. We are also interested in exploring the usefulness of non-explicit communication through affective channels [22]. Unobtrusive sensing of temperature or skin conductivity, displayed graphically, might add a valuable human element to individuals' representation.

Another view of the system is that of Benford et al's schema for shared spaces. According to their criteria *Talking in Circles* is of medium transportation and artificiality, and of extremely high spatiality, supporting on-going activity, peripheral awareness, navigation and chance encounters, usability through natural metaphors, and a shared frame of reference [2]. We are interested in extending the system's spatiality even further, such as by providing greater persistence and meaning to the space, that is, increasing its sense of place [9]. This might be done by modifying the background of different chat sessions or by permitting wear on the space, such as permanent user drawings or subtly showing which areas of the space get the most use. We are also interested in extending our framework for participants' navigation along architectural notions. The social significance of central and peripheral areas in plazas or the flow of people at street corners [33] indicates potential for interface techniques to make our space more legible and navigable.

A related area for future work is in preserving history in an audio chat. *ChatCircles* has used spatially-useful history mechanisms based on conversation groups at various points during a chat [32], but parsing and browsing of sound remains a major challenge. Braided audio is one interesting recent approach [26].

A suggestion we have heard often is making the audio fully 2D-spatialized rather than the current 1D attenuation. This could certainly add to the system's spatiality, but we first intend to observe how straightforward users' mapping of spatialized audio is onto the flat 2D surface *Talking in Circles* presents. A mechanism for explicitly conveying gaze direction, such as a more abstract analogue to Donath's work with pictures of chat participants facing in different directions [6], might also be of great benefit.

REFERENCES

1. Ackerman, M., Hindus, D., Mainwaring, S., Starr, B. Hanging on the 'wire: A field study of an audio-only media space. In *ACM Transactions on Computer-Human Interaction*, March 1997, 4(1), pp. 39-66.
2. Benford, S., Brown, C., Reynard, G., Greenhalgh, C. Shared Spaces: Transportation, Artificiality and Spatiality. In *Proceedings of CSCW '96*, ACM, 1996, pp. 77-86.
3. Chalfonte, B., Fish, R. S., Kraut, R. E. Expressive Richness: A Comparison of Speech and Text As Media for Revision. In *Proceedings of CHI '91*, ACM, 1991, pp. 21-26.
4. Chapanis, A. Interactive Human Communication. *Scientific American*. 232, 1975, pp. 36-42.

5. Cherry, E. C. Some experiments on the recognition of speech with one and with two ears. *Journal of the Acoustical Society of America*, 25, 1953, pp. 975-979.
6. Donath, J. The Illustrated Conversation. *Multimedia Tools and Applications*, 1, March 1995.
7. Erickson, T. et al. Socially Translucent Systems: Social Proxies, Persistent Conversation, and the Design of "Babble." In *Proceedings of CHI '99*, ACM, 1999, pp. 72-79.
8. Greenhalgh, C., Benford, S. MASSIVE: a collaborative virual environment for teleconferencing. *ACM Transactions on Computer-Human Interaction.* 2(3), 1995.
9. Harrison, S., Dourish, P. Re-Place-ing Space: The Roles of Place and Space in Collaborative Systems. In *Proceedings of CSCW '96*, ACM, 1996, pp. 67-76.
10. Hindus, D., Ackerman, M. S., Mainwaring, S., Starr, B. Thunderwire: A Field Study of an Audio-Only Media Space. In *Proceedings of CSCW '96*, ACM, 1996, pp. 238-247.
11. Isaacs, E. A., Tang, J. C. What Video Can and Can't Do for Collaboration: a Case Study. In *Proceedings of Multimedia '93*, ACM, 1993, pp. 199-206.
12. Java Sound API. Sun Microsystems, Inc. http://java.sun.com/products/java-media/sound
13. Kansas. Sun Microsystems, Inc. http://research.sun.com/research/ics/kansas.html
14. Krauss, R. M., Bricker, P.D. Effects of transmission delay and access delay on the efficiency of verbal communication. *Journal of the Acoustic Society of America*. 41, 1967, pp. 286-292.
15. Kroll, B.M. Cognitive egocentrism and the problem of audience awareness in written discourse. *Research in the Teaching of English*. 12, 1978, pp. 269-281.
16. Kurlander, D., Skelly, T., Salesin, D. Comic Chat. In *Proceedings of SIGGRAPH*, 1996, pp. 225-236.
17. McDaniel, S. E., Olson, G. M., Magee, J. C. Identifying and Analyzing Multiple Threads in Computer-Mediated and Face-to-Face Conversations. In *Proceedings of CSCW '96*, ACM, 1996, pp. 39-47.
18. Milgram, S.. *The Individual in a Social World.* Reading, MA: Addison-Wesley, 1977.
19. Nakanishi, H., Yoshida, C., Nishimura, T., Ishida, T. FreeWalk: Supporting Casual Meetings in a Network. In *Proceedings of CSCW '96*, ACM, 1996, pp. 308-314.
20. O'Conaill, B., Whittaker, S., Wilbur, S. Conversations Over Video Conferences: An Evaluation fo the Spoken Aspects of Video-Mediated Communication. *Human-Computer Interaction.* 8, 1993, pp. 389-428.
21. The Palace, Inc. The Palace. http://www.thepalace.com
22. Picard, R. *Affective Computing*. Cambridge, MA: MIT Press, 1997.
23. Progressive Networks, Inc., http://www.real.com
24. Rutter, D.R. The role of cuelessness in social interaction: An examination of teaching by telephone. In Roger, D. and P. Bull (ed.), *Conversation: An Interdisciplinary Perspective*. Multilingual Matters, 1989.
25. Savetz, K., Randall, N., Lepage, Y. *MBONE: Multicasting Tomorrow's Internet*. IDG, 1996.
26. Schmandt, C. Audio Hallway: a Virtual Acoustic Environment for Browsing. In *Proceedings of UIST '98*, San Francisco, CA: ACM, 1998, pp. 163-170.
27. Sellen, A. J. Speech Patterns In Video-Mediated Conversations. In *Proceedings of CHI '92*, ACM, 1992, pp. 49-59.
28. Singer, A., Hindus, D., Stifelman, L., White, S. Tangible Progress: Less Is More In Somewire Audio Spaces. In *Proceedings of CHI '99*, ACM, 1999, pp. 104-111.
29. Smith, G. Cooperative Virtual Environments: lessons from 2D multi user interfaces. In *Proceedings of CSCW '96*, ACM, 1996, pp. 390-398.
30. Tang, J.C. Findings from observational studies of collaborative work. *International Journal of Man-Machine Studies*. 34(2), 1991, pp. 143-160.
31. University College London. Robust Audio Tool. http://www-mice.cs.ucl.ac.uk/multimedia/software/rat
32. Viegas, F. B., Donath, J. S. *ChatCircles*. In *Proceedings of CHI '99*, ACM, 1999, pp. 9-16.
33. Whyte, W. H. *City: Rediscovering the Center*. NY: Doubleday, 1988

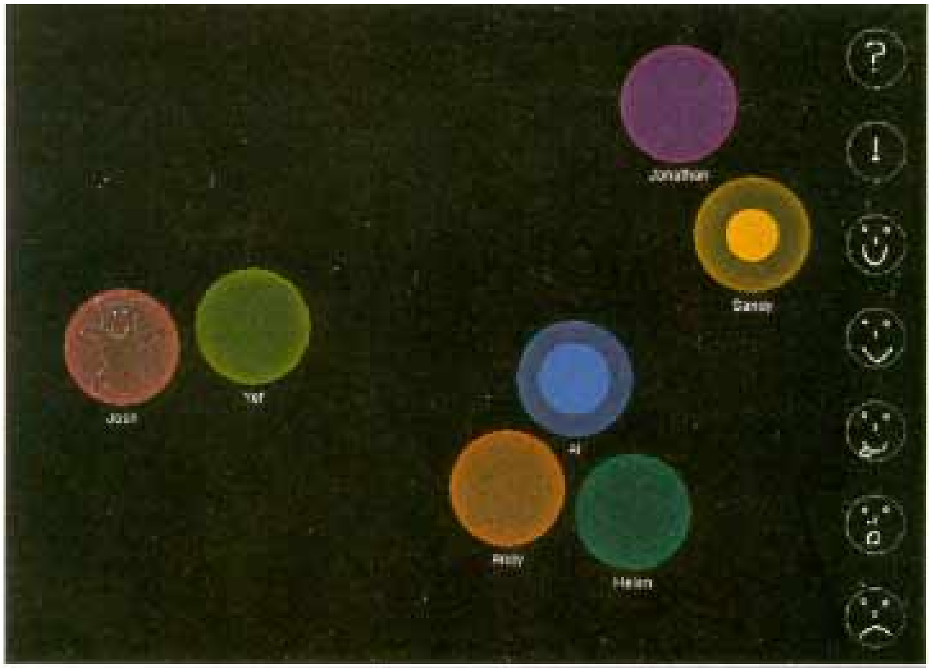

Figure 1: Three conversations in a *Talking in Circles* chat session. Al is talking with Helen and Andy, and Sandy converses with Jonathan. Meanwhile, Josh draws for Yef.

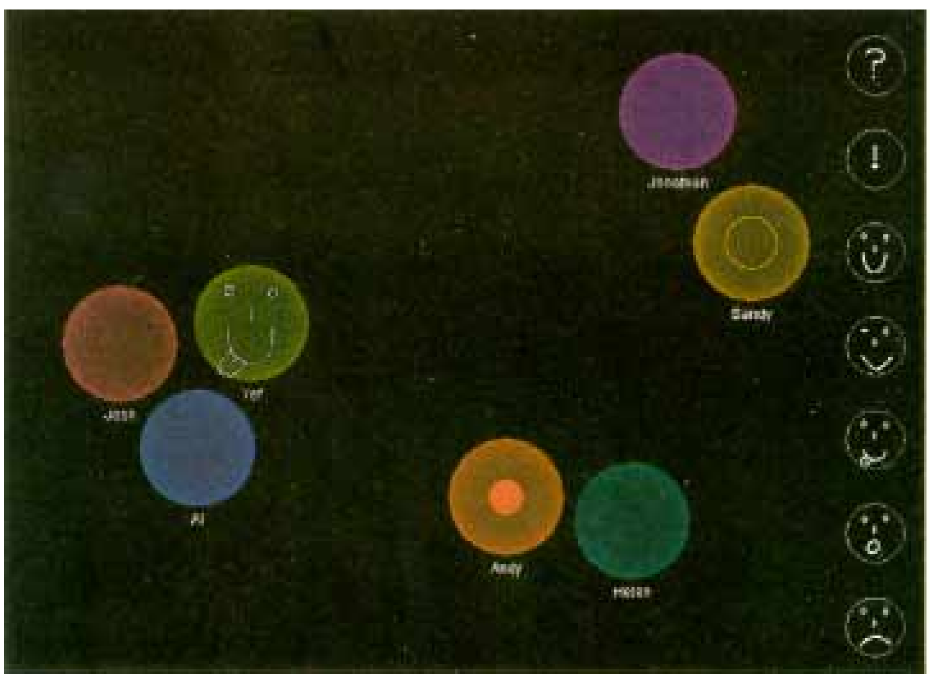

Figure 2: Screenshot from the point of view of Al, the blue circle, after he has wandered over to speak with Josh and Yef. Both Sandy and Andy are speaking, but Andy remains within Al's hearing range while Sandy is now outside it.