

# Approaching and leave-taking: Negotiating contact in computer-mediated communication

JOHN C. TANG

Sun Microsystems Laboratories

---

A crucial difference between face-to-face interaction and computer-mediated communication is how contact negotiation—the way in which people start and end conversations—is managed. Contact negotiation is especially problematic for distributed group members who are separated by distance and thus do not share many of the cues needed to help mediate interaction. An understanding from conversation analysis of what resources and cues people use to negotiate making contact identifies ways to design support for contact negotiation in new technology to support remote collaboration. This perspective is used to analyze the design and use experiences with three communication prototypes: The Desktop Conferencing Prototype, Montage, and Awarenex. These prototypes use text, video, and graphic indicators to share the cues needed to gracefully start and end conversations. Reviewing what we learned from these research experiences identifies directions for future research in supporting contact negotiation in computer-mediated communication.

Categories and Subject Descriptors: H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces – *Computer-supported cooperative work; Synchronous interaction, Collaborative computing*; H.4.3 [**Information Systems Applications**]: Communications Applications – *Computer conferencing*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *User-centered design*

General Terms: Design, Human Factors

Additional Key Words and Phrases: Computer-mediated communication, human-computer interaction, interaction design, user research

---

Author's address: John C. Tang, Sun Microsystems Laboratories, 2600 Casey Avenue, MTV29-230, Mountain View, CA 94043 USA, John.Tang@sun.com.

Permission to make digital/hard copy of part of this work for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication, and its date of appear, and notice is given that copying is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee.

© 2001 ACM 1073-0516/01/0300-0034 \$5.00

## 1. DESIGNING COMPUTER-MEDIATED COMMUNICATION TO SUPPORT DISTRIBUTED TEAMS

Starting and ending conversations in computer-mediated communication media is more clumsy than doing so face-to-face. The effort in answering telephones, starting video conferences, responding to instant messages (IMs), replying to e-mail, etc. contrasts with the apparent ease in starting and ending conversations when face-to-face. This observation demonstrates both how accomplished people are in gracefully negotiating everyday conversation, and how impoverished technology-mediated channels are in supporting that negotiation.

Supporting this process of starting and ending contact, which I will refer to as *contact negotiation*, is especially important for distributed groups who rely on communication technology to help them establish contact. Without the cues that face-to-face interaction affords, remote collaborators depend on the support for contact negotiation that the design of the communication technology offers. This design problem has led to a line of research that has applied an understanding of how people negotiate starting and ending conversations to the design of new communication technologies.

In this paper, I review the interaction design of three research prototypes to support communication among remote collaborators: The Desktop Conferencing Prototype, Montage, and Awarenex. I will describe our design intentions for each prototype and review what we learned from studying each prototype in actual use. Reflecting on our experiences with these prototypes documents a progression of design ideas that suggests new research directions for how to better support contact negotiation in computer-mediated communication.

### 1.1 The Nature of Distributed Work

The ongoing development of new technologies to support communication among remote participants has been concurrent with an increasing willingness to distribute work across different locations. E-mail, mobile telephony, group chat, IM, and desktop video conferencing, have come into broad use among distributed teams. At the same time, trends in work environments have created groups that include people distributed among different sites, telecommuters from home, and collaborations that span the globe. While everyday work practice has come to rely on communication tools (e.g., e-mail, phone, IM) to help contact the remote people needed to accomplish work, interacting through those tools often feels less satisfying than conversing face-to-face.

Prior research on informal, face-to-face interactions has identified contact negotiation as an important focus of attention. Kraut et al. [1990] described how physical proximity

affords frequent, low-cost, high fidelity communication (i.e., ease of contact negotiation) that enables developing satisfying working relationships. Whittaker et al. [1994] showed how co-located collaborators drew upon contextual cues to be aware of relevant events and negotiate when to interrupt others to establish contact.

Studies have compared face-to-face with computer-mediated communication along various aspects: social psychological [Kiesler et al. 1984], group dynamics [Siegel et al. 1986], and efficiency [Fish 1988]. Whittaker's [in press] thorough review of theoretical approaches to comparing face-to-face and computer-mediated communication grouped prior research according to whether it focused on:

- the raw sensory information
- the cognitive cues that a person uses to interpret that sensing information into behaviors
- the social cues that people use together to provide a context for interpreting the behaviors and information

Whittaker called for going beyond sharing the raw sensory information to further study how computer-mediated communication affects the cognitive and social cues in interaction. Focusing on sharing the cognitive and social cues that people use to mediate their communication does not require literally re-creating the sensory inputs that provide those cues. This paper applies this perspective to consider the ways people use sensory, cognitive, and social cues to accomplish contact negotiation and explore new ways to support it.

## 1.2 Designing to Support Contact Negotiation

Face-to-face contact negotiation relies on many subtle and non-verbal cues (e.g., eye gaze, head nods, gestures). Several research projects (e.g., [Bly et al., 1993; Dourish & Bly, 1992; Fish et al., 1992]) have explored using video to share visual and aural cues across distance to support contact negotiation. Recent research [Cassell & Vilhjálmsón 1999] continues the tradition of trying to re-create face-to-face sensory cues in mediated conversations through a virtual avatar.

Figuring out how to accommodate the diminished resources available for remote participants is a design challenge for computer-mediated communication. The Coordinator [Flores et al., 1988] was an early attempt to apply computer technology to augment the communication process among people. The ensuing debate [Suchman, 1994; Bannon, 1995] illustrated the potential problems with trying to fit the richness of human communication into cleanly defined categories for action. Buxton [1997]

summarizes more recent research efforts to combine video and sensor information to mediate approaching contact with remote collaborators.

This paper draws on understanding (mainly from conversation analysis) how people accomplish starting and ending conversations. That understanding is used to analyze the design and use experiences with several computer-mediated communication systems and prototypes. This analysis synthesizes what we have learned from these research experiences about designing computer-mediated communication and identifies areas for future research.

## 2. DRAWING ON CONVERSATION ANALYSIS

The field of conversation analysis has been studying human communication in a variety of situations, including face-to-face [Sacks et al. 1974; Heritage 1984] and telephone [Schegloff 1972] interactions, to understand how people accomplish conversation. Conversation analysis researchers have identified many of the cues and resources people use to mediate turn-taking, organize the structure of talk, and start and end conversations.

In conversation analysis terminology, the beginnings of conversations are called *openings* and the endings are *closings*. Taken together, openings and closings form the focus of this paper on contact negotiation. Studies of openings and closings go as far back as Goffman's [1963] research in studying encounters in public places. Kendon and Ferber [1973] looked at openings in face-to-face interaction and Knapp et al. [1973] looked at closings in face-to-face interaction. These studies can be compared with the analysis of openings and closings drawn largely from telephone conversations [Schegloff 1972; 1973], which are distinctly different than face-to-face interactions. Clark [1985] summarizes the studies of openings and closings in the following structural framework.

Openings consist of:

- *contact initiation*—mutually recognizing an attempt to initiate contact
- *greetings*—establishing each person's identity and that a conversation has started
- *topic initiation*—introducing the first topic

Closings consist of:

- *topic termination*—mutually recognizing that the topic discussion has ended
- *leave-taking*—reaffirming each other's acquaintance before breaking contact
- *contact termination*—ending the connection that was enabling the conversation

Clark explains leave-taking as a process where, "...people often summarize the content of the contact they have just had; justify ending their contact at this time; express pleasure about each other; indicate continuity in their relationship by planning, specifically or vaguely, for future contact; and wish each other well..." Leave-taking

smoothes the transition from expressing intent to end the contact to actually disconnecting from the conversation, allowing time to affirm with the others that it is an agreeable time to end the conversation.

While there is a wide range of ways to accomplish openings and closings, the functions represented in these steps should be fulfilled (even if most minimally) in all interactions. For example, people who are familiar with each other may tacitly accomplish greetings through recognizing each other by sight (or by voice over the phone). Similarly, leave-taking is tacitly accomplished by a pause in the conversation and starting to move away from each other.

Just as important as successfully accomplishing conversation openings is gracefully avoiding an opening if now is not a time to have a conversation. As Goffman [19663] points out, opening moves in face-to-face interaction are often done subtly and tentatively (e.g., a quick eye glance), so that one can still act as if no initiation was intended if the recipient does not want to engage in conversation. It is this gracefully negotiated avoidance of contact where current communication technologies are most noticeably impoverished compared to face-to-face interaction.

To gain familiarity with Clark's framework, it will be used to first review how openings and closings are accomplished in face-to-face interaction and telephone calls. Then, this framework will be used to explain how we designed three prototype tools to support communication among remote participants.

### 3. CONTACT NEGOTIATION IN FAMILIAR CONTEXTS

#### 3.1 Face-to-face interaction

Contact initiation is accomplished by encountering each other, establishing mutual eye contact, positioning bodies to enable direct conversation, perhaps pausing, and perhaps using gestures and other nonverbal cues to start a conversation. Many of these contact initiation moves are offered tentatively so that they can be gracefully overlooked if no interaction is desired. Face-to-face greetings afford identifying each other tacitly by visual or aural recognition. Face-to-face greetings also include verbal exchanges that confirm that it is an appropriate time to have a conversation, and often set the context for availability ("I only have a couple minutes").

Topic termination often naturally occurs when the topic of conversation resolves, often accompanied by pauses, closure statements ("I think that's it"), or nonverbal cues of being satisfied or complete (shifting posture, nodding head). Leave-taking exchanges often shift to the interpersonal relationship (expressing thanks, anticipating the next

opportunity to meet). These verbal exchanges are often accompanied by nonverbal actions indicating the intent to end the conversation (standing up from a chair, stowing away books, laptop computers, and any props used in the conversation). Furthermore, this leave-taking process can be aborted at any time if one of the participants has more to say before ending the conversation. As the correspondents move out of hearing range from each other, contact termination is completed.

In informal, face-to-face conversations, Whittaker et al. [1994] observed that the greeting and leave-taking steps are extremely abbreviated or even skipped over, especially if this is not their first encounter of the day. Thus, they characterized repeated informal contacts throughout the day as a conversation that gets suspended and resumed almost without any openings and closings in the subsequent encounters.

### 3.2 Telephone conversations

While telephone conversations have become as familiar as face-to-face interactions, the telephone mediates conversation in fundamentally different ways. The framework for openings and closings will highlight how the affordances of telephones affect how much information is shared between the participants and how people have accommodated to the phone.

In telephone calls, the caller must start contact initiation by dialing a phone number. On the recipient's side, contact initiation begins with the phone ringing. At this point, the recipient needs to make a decision about whether to interrupt her current activity to answer the call. The recipient has very little information about the incoming call (who is calling? how urgent is this issue?), whereas in face-to-face interaction, she would know who is approaching and may detect some nonverbal cues about the topic. While caller-id may indicate who the caller is, this is not yet in pervasive use (and is sometimes fallible, as the caller's identity is associated with the originating phone number, not the caller). Also, unlike the subtle contact initiation moves when face-to-face, it is more difficult for the recipient to ignore a ringing telephone.

Contact initiation is completed when the recipient decides to pick up the phone. Greetings on the phone begin with routine exchanges of "hello" (confirming that everyone can hear each other) and identifying each other (either by stating names or recognizing voices). From here, the rest of greeting and topic initiation proceeds similarly to face-to-face interactions, using only the aural cues from the phone.

While topic termination on the telephone also tends to sound very similar to that in face-to-face, leave-taking is typically protracted to explicitly confirm that all parties agree

that the conversation is completed (“That’s all I have”) followed by at least one round of everyone explicitly saying “goodbye”. This extra confirmation helps avoid prematurely hanging up the phone and the overhead of re-dialing after if contact was terminated too soon.

#### 4. DESIGNING CONTACT NEGOTIATION IN PROTOTYPE COMMUNICATION TOOLS

Understanding the cues used in contact negotiation helps define what cues need to be shared through computer-mediated communication tools. Even the contrasts between contact negotiation in face-to-face and telephone interactions are instructive. One important attribute of attempts to initiate contact initiation is how gracefully they can be avoided or deferred to another time. Another important question to consider is how much information does a person have before committing to participate in an interaction? In face-to-face conversation, one knows who is approaching and may have some topic cues *before* the conversation starts. They also share some cues about availability to start an interaction. On the phone, however, the recipient must decide whether to pick up the phone *without knowing* who is calling or any topic context. Furthermore, the caller has no context of the recipient's availability.

The remainder of this paper describes the design of three communication prototypes that take advantage of capabilities among networked computers to support the communication and coordination of teams who are distributed among different locations:

- DCP — an early Desktop Conferencing Prototype
- Montage — a prototype interface for initiating desktop video conferencing
- Awarenex — an IM and awareness prototype

For each prototype, I describe the design intent for supporting contact negotiation and the lessons learned from our experiences using the prototype. Our design approach follows the tradition of social translucence [Erickson & Kellogg 2000] by providing the participants the cues that allow them to socially negotiate appropriate action among themselves. Building on the existing understanding of contact negotiation, we can evaluate how the prototype designs support the ability for smoothly negotiating contact.

##### 4.1 Desktop Conferencing Prototype (DCP)

The Desktop Conferencing Prototype (DCP) was designed in 1991, when the capability of real-time audio and video connections among computer desktops was just emerging [Tang & Isaacs 1993]. Figure 1 shows the interface for setting up desktop conferences

that enables users to select who they wanted to conference with and what features (audio, video, shared whiteboard) to launch.

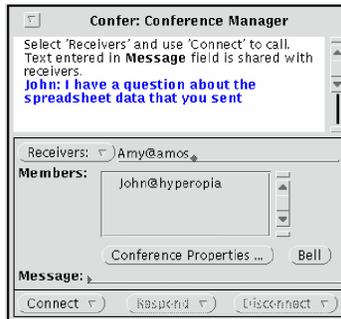


Figure 1: Interface for initiating a desktop video conference with a shared text area for negotiating making contact.

Since it took a significant amount of time (about a half minute) to make all the audio and video connections among desktops, we wanted to provide an interface that helped users determine if it was a good time to initiate a conference before actually going through the trouble of doing so. We integrated a chat-like shared message area in the interface that allowed prospective participants to exchange text messages to confirm if they were available. We wanted to give users a means for negotiating contact initiation and greetings (and perhaps agree on topic initiation) to confirm it was worth the effort and time needed to start a conference.

Besides using the DCP among our own group (split between East and West Coast sites in the U.S.), we deployed it in a team of five people distributed over one East Coast and two West Coast work sites. These users were part of a product team that was not related to our research group. We studied their use of the DCP for seven weeks [Tang & Isaacs 1993]. Besides logging statistics of usage, we interviewed the team to gather their perceptions on the system. From their usage, we learned how the interface supported their contact negotiation.

One main finding from the DCP study is that only 43% of the attempts to initiate conferencing resulted in a desktop conference. Many contact attempts did not result in a conference because the recipient was not in the office or not available to respond to the request. This finding suggested the importance of supporting the *pre-interaction negotiation* that allows collaborators to find opportune times to contact each other. Especially since the overhead in initiating contact was so high, it would have been helpful to have some cues to help determine whether attempting to make contact was worth the effort.

While some users appreciated the design of the shared message area in the interface, we discovered at least one user who did not bother to use it. He explained that it took too long (over ten seconds) to launch the interface program which was used to initiate conferences. He found it quicker to call people on the phone (while waiting for the interface to come up), to ask if they were available. Thus, he would use the phone, rather than the shared message area designed in the interface, to accomplish the contact initiation and greeting functions, and simply used the interface to launch the conferencing connections. Supporting contact negotiation needs to be at least as quick and lightweight as other comparable technologies, such as the phone.

For conversation closings, once the participants agreed that the conference was over, each person selected the “Disconnect” button on the interface, which closed all associated conferencing windows (e.g., video, shared whiteboard, audio connection). In this sense, ending desktop conferences was similar to ending telephone conversations, with protracted leave-taking where each participant explicitly confirmed that they were done, usually followed by each person saying “goodbye”. This comprehensive closing process was needed to avoid the substantial effort that premature contact termination would incur.

## 4.2 Montage

Our experience with the DCP led us to investigate better ways of supporting contact initiation. In 1993, we designed the Montage interface to explore the use of video to support contact initiation, rather than the text messages used in the DCP [Tang & Rua 1994]. This approach builds on the ways contact initiation is typically done in face-to-face settings where people typically visually assess whether a person is available for a conversation. Establishing mutual eye contact is often the first cue of being available and accomplishing contact initiation. In Montage, we gave users a video “glance” between them to help assess whether they were available.

A video-only glance immediately shows if the intended recipient is not in the office or unavailable (on the phone, with visitors). Visual glances are especially useful to support contact initiation since many of those cues are non-verbal (eye gaze, hand wave, head nod). A visual channel addresses many of the problems that we saw using the DCP and enables remote collaborators to accomplish contact initiation and greeting in ways similar to face-to-face openings.

Of course, opening a video glance with another computer user could be viewed as distracting and raise privacy concerns, as discovered with the Cruiser prototype at Bellcore [Fish et al. 1992]. In the Montage design, shown in Figure 2, the video glance

window gradually faded in (accompanied with an audio sound) to provide a sense of approach. Just as people do not suddenly appear in your office in face-to-face encounters, we provided a similar sense of approach for video glances that appear on your computer desktop.

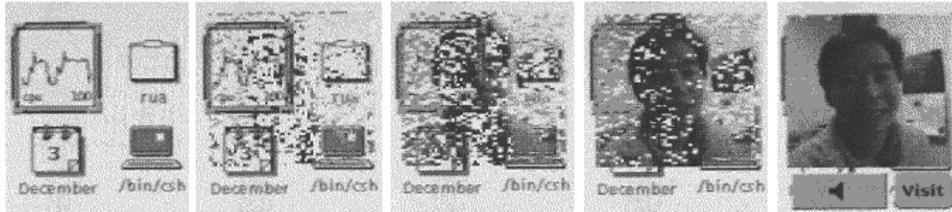


Figure 2: Sequence of images (in black and white) showing how a video glance fades in to provide a sense of approach.

Video glances were reciprocal; at the same time that a video glance of you faded in on my desktop, a video glance of me faded in on your desktop. They were also momentary; if neither person reacted to the glance, it faded out after eight seconds. Thus, the glance is modeled after how someone might walk by your office to see if you are available. You would become increasingly aware of them as they approached, and if you did not react to them or they could visually see that you were unavailable, they would walk past without disturbing you.

Montage glances were also momentary; a video-only glance fades in, holds for eight seconds, then fades out if no one reacts to it. This design made avoiding contact initiation the default action (glance fades out without a response), with the intent that such avoidance would not appear rude. Joining a glance was easily accomplished by clicking a button. Based on our DCP experience, we also architected Montage to “pre-start” the video and audio connections, so that selecting a person started a glance almost immediately (less than three seconds) to maintain a lightweight feel when glancing.

If the video glance suggested that now would be a good time to interact, either party could acknowledge the glance by pressing on the audio button. This turned on the audio channel, effectively starting a desktop video conference. At this point, the video windows were relatively small (128 x 120 pixels), but for extended interactions, participants could press the “Visit” button to enlarge the video windows (256 x 240 pixels).

The interface for the person initiating the glance is shown in Figure 3. Besides the glance interface, it also offered quick access to on-line calendar, e-mail, and an electronic note tool that could help coordinate making contact in the future. If I glance at your office and find that you are not there, I could bring up your on-line calendar to see if you have

an appointment that accounts for your absence (and suggests when you might return). I could send e-mail, transforming a synchronous contact attempt into an asynchronous one. Or, I could leave an electronic note on your screen that could explain my intent in contacting you. The note has a “Glance Back” button embedded in it, making it easy to respond to the note by glancing back at me.



Figure 3: Glance window with the audio link enabled. The window also affords easy access to on-line calendar, e-mail, and electronic note tools to negotiate future contact.

The Montage interface provided many cues and tools needed to negotiate smoothly starting an interaction. With Montage, users were able to accomplish contact initiation and greetings much as in face-to-face interaction. They were able to see who was approaching and visual cues about availability and urgency of topic. As in face-to-face interaction, this pre-interaction context was shared in Montage before committing to the interaction. If it was not a good time to make contact, the glance faded out gracefully, but if it was a good time to interact, they could easily initiate desktop conferencing. Again, contact termination was accomplished by clicking a button that closed all video windows and audio connections.

Besides using Montage among our own group (which was at times distributed among a couple buildings on a campus), we also deployed Montage among ten people in an unrelated product group who were distributed among three buildings on a campus [Tang et al. 1994]. We studied their use of Montage for twelve weeks. As with our study of the DCP, we recorded many more attempts to contact someone than glances that turned into desktop conferences. Overall, only 25% of glances turned into desktop conferences. This conversion rate is actually less than that of the DCP (43%). Despite this low conversion rate, we found it encouraging that users continued to use Montage throughout the duration of the study. Furthermore, comments in the interviews indicated that they liked Montage because it was lightweight and easy enough to use, even though it often did not result in a desktop conference. Even if a glance showed that the person was unavailable for contact at the moment, the interface provided opportunities to coordinate future interactions through the calendar and electronic note tools.

In our videotaped studies of Montage use, we focused on openings and closings by transcribing and analyzing all instances (both using Montage and accomplished face-to-face) captured on videotape. We found that openings in Montage were similar to face-to-face openings in that nonverbal gestures and cues were used for contact initiation and greeting. We observed gestures that indicated needing to wait a minute, being currently unavailable but would return contact soon, or being occupied on the telephone. We also saw examples of interpreting visual information (e.g., noticing a user still wearing a bike helmet, indicating that he had just gotten in and was rushed).

Closings in Montage, on the other hand, were more like telephone interactions, where an explicit round of goodbyes was exchanged before pressing the button that closed all the video windows and disconnected the audio. While the video gave them more cues than would be available on the phone, the interface did not provide any support for leave-taking to disconnect the windows. We even observed an example where one user disconnected the conference just as the other user was about to add a last-minute thought, and they had to go through the overhead of re-establishing the connection to complete the conversation. While we appreciated the ways we supported openings through the interface, our experience with Montage drew our attention to better support for closings.

### 4.3 Awarenex

The Awarenex prototype [Tang et al. 2001], developed in 2000, represents a different approach to supporting distributed communication. Rather than including a video channel, Awarenex relies on textual and iconic representations of communication (IM, telephone) and awareness information. This approach was in part due to the lack of penetration of a networked video infrastructure, contrasted with the popularity of IM systems. It also relies less on re-creating the sensory inputs from face-to-face interaction and more on providing the information needed to make the cognitive and social inferences for negotiating contact.

Awarenex consists of a Contact List of people of whom you want to maintain awareness. As shown in Figure 4, each entry in the Contact List displays:

- name
- locale (e.g., home, office, lab, mobile)
- any keyboard idle time
- activity indicators (in an IM, on the phone, in an appointment scheduled in an on-line calendar)



Figure 4: Awarenex Contact List showing name, locale, keyboard idle time, and any activity indicators for each entry.

As with commercial IM systems, the awareness information presented in the Contact List provides an initial clue about availability for contact. Double-clicking on any entry in the Contact List initiates an attempt to establish an IM with that person. Based on our experiences with the DCP and Montage, we designed the interface for starting and ending IMs to support the opening and closing processes.

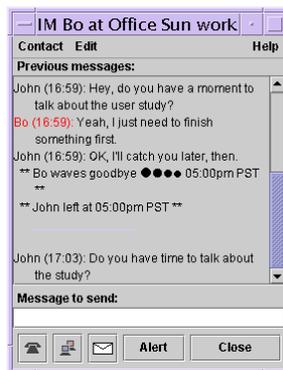


Figure 5: Initiating an IM, which starts out with a gray background until the recipient joins the IM session.

Previous IM sessions are available by scrolling backward.

Figure 5 shows the IM window that initially comes up when I select to IM Bo. The background color of the text area is initially gray until Bo joins the IM, when it turns to white. This indicator of mutual attention addresses the problem in many IM systems where the first exchanges of conversation often confirm whether the person is really “there” and attending to the IM.

The IM window also contains the transcript of previous IMs that I have had with Bo earlier in the day. This design reflects the findings from Whittaker et al. [1994] that informal communication often feels like a sustained but interrupted conversation that gets paused and restarted throughout that day. The IM session contains the text captured from previous IMs between those people to provide prior context for the current IM. That transcript gets cleared at the end of each day.

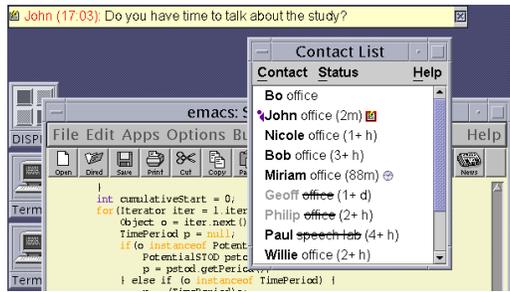


Figure 6: Contact Preview of an incoming IM that rolls down from the top edge of the computer screen.

As the recipient, Bo first gets a Contact Preview from me, shown in Figure 6. The Contact Preview slides down from the top of Bo's screen, accompanied with a sound alert, that shows the first line of text that I am typing in the IM. The Contact Preview gives a subtle sense of approach, and gives an indication of topic initiation of what the IM is about. This contrasts with most commercial IMs that immediately pop up a window on the recipient's computer screen without any sense of approach. Bo can join the IM by clicking on the Contact Preview, or dismiss the Contact Preview by clicking the close box. Or, Bo could simply disregard the Contact Preview, leaving it there until he decides to join the IM at a later time or the initiator withdraws the IM request.

Once Bo joins the IM, turning my IM text background to white, we can engage in an IM conversation. We transmit the text input character-by-character to support more of a sense of conversational flow in IM. Seeing the other person's typing appear character-by-character not only provides a cue of whether that person is in the process of replying (instead of attending to other things at the time), but also affords the possibility of anticipating what is being said before actually completing all the typing.

Awarenex supports contact initiation in ways that are different than the previous video-based prototypes but similar to IM systems. While IM users do not share a rich visual channel, they share some awareness cues to suggest whether they are available for interaction. In their study of IM usage, Nardi et al. [2000] found that this awareness information and a lightweight IM exchange were often used to establish if a person was available for contact. If so, the conversation often migrated to a phone call, as a more efficient medium of communication. They also found that people could ignore incoming IMs without offending the sender because there was enough ambiguity in the awareness information that people realized there could be a number of plausible reasons for not responding to an IM. Our usage experiences with Awarenex were similar, since it shared these design attributes with IM.

The design of Awarenex also supports contact negotiation in some new ways beyond typical IM systems. In Awarenex, the recipient of an IM request actually knows who is

making the request and some sense of topic initiation before committing to respond to the IM. The Contact Preview is often used to convey what the IM request is about, similar to the way the subject field in an e-mail message is used. This interface design offers a novel affordance of identifying topic initiation before committing to an interaction. Also, since IMs often migrated into phone calls and more extensive collaboration, we integrated buttons in the Awarenex IM that connected a phone call and launched application sharing between the IM participants.

Figure 7 shows the leave-taking interface designed in Awarenex. In contrast to most commercial IM systems, which close the windows as soon as a user ends the IM, Awarenex has a “Goodbye” button that initiates the leave-taking interface. Pressing “Goodbye” prints a system generated message in the interface (“John waves goodbye”) and starts a series of diminishing dots. This indicates my intent to end the IM, but leaves the IM connection open to provide some time (ten seconds) to negotiate any last-minute messages. This leave-taking process is modeled after closings in face-to-face interaction, where participants often give nonverbal cues that they are ready to leave (e.g., close notebook, stow pens, pick up briefcase) while still allowing time to gracefully negotiate ending the conversation. Users can abort the leave-taking process by simply typing into the IM or pressing the “Rejoin” button. If they let the ten-second countdown expire, the IM window closes.



Figure 7: IM window showing that John has initiated the leave-taking interface, indicated by the series of diminishing dots, while Bo wants to express a last-minute change in plans.

While we have not formally studied the use of Awarenex, about 50 people have used it over the course of two years. These users have included teams distributed between East and West Coasts in the U.S. and product teams unrelated to our research group. While the awareness information presented by text and icons is less rich than the visual information conveyed by video in our earlier prototypes, the ease of installation makes

the use of Awarenex pervasively available (e.g., at home, on a mobile device). The text and icons can efficiently identify people and convey some contextual information about their activities in ways that help mediate communication. Isaacs, et al. [2002] also explored using sound cues to help identify people's activities and conversations without having to look directly at the interface. In these ways, the functions of contact initiation, greeting, and topic initiation can be accomplished without actually re-creating the gestures and other non-verbal cues used in face-to-face interaction.

Based on our use experience with Awarenex, we found that sharing the awareness information displayed in the Contact List is valuable among a distributed team, even without much use of IM. Logs of six of the longest running Awarenex users indicate that the average number of IMs that each user sent ranged from only 0.67 to 2.21 messages per day. Awareness of when people are at their computer is useful in timing when to place a phone call to them. This awareness is especially useful for people in different time zones. East Coast users can get a sense of when West Coast colleagues start their days in the office. Both sites can coordinate activities with when they leave and return from lunch, and the West Coast users know when East Coast colleagues leave at the end of the day (and if they check e-mail from home in the evening before the end of the day on the West Coast). Integrating other coordinating cues, such as whether they have an appointment scheduled in their on-line calendar, also helps identify good times to make contact.

Many people appreciate the Contact Preview interface, although some prefer having the IM window open automatically, based on familiarity with commercial IM systems. We have made this approach preference a user configurable setting. Users often mention liking the leave-taking interface. Although we did not log instances of aborting the leave-taking process, we have observed instances of doing so in IM interactions, so believe that the design does afford negotiating last-minute interruptions before closing the windows.

## 5. THE FUTURE OF COMMUNICATION INTERFACES

As more and more of our communication becomes mediated by technology (e.g., e-mail, mobile phones, mobile text messaging), we need to focus on designing interfaces that support human communication practices and their environments of use. Especially as communication services migrate to computer-mediated and other new technology infrastructures, opportunities to design interfaces that provide better contact negotiation support should be explored. Unfortunately, much of the current work in this area simply

recreates existing capabilities on computer platforms. For example, the Voice over IP (VoIP) work to date has largely focused on re-creating telephony interfaces on a computer-based, internet packet network, without taking advantage of integrating other computational capabilities that could support the communication process. By reviewing the design and user experiences with our research prototypes, I hope to have identified points of departure for future interface design.

Some interfaces entering into widespread public use have evolved more from what technology can do rather than an understanding of how people communicate. For example, the traditional phone interface of ringing the phone of a recipient was designed for placing a call to phones in fixed locations where nearby people needed to be alerted of an incoming call. But now, phones have become mobile devices that individuals carry with them into a variety of contexts. The sound of ringing phones and the ensuing conversations in public places has the potential of disrupting a large number of people.

A related observation demonstrates how conversations while driving a car are different between talking with a passenger versus talking over a mobile phone. Research by Parkes [1991] showed measures indicating that driving while talking over the phone is more difficult than talking with a passenger. This difference is attributed to the shared context that the car passenger has that helps make allowances for traffic conditions while conversing.

A simple way to re-examine the phone interface in a mobile setting is to change how incoming calls are announced. Thus, recent mobile phones typically offer a “vibrate” mode which is less obtrusive. Recent research has explored alternative interfaces for handling mobile phone calls in public settings that provide some aural [Nelson et al. 2001] or textual [Pedersen 2001] indication of topic initiation of the call before fully committing to it. A fundamental problem for further research is that callers are currently totally blind to the context of the recipient, and thus have no cues whether this would be an appropriate time and place to initiate a call. Finding ways to share the cues among callers, recipients, and their surrounding environments may help create more graceful interfaces for initiating and managing calls among mobile phone users in public places.

As computer technology mediates more and more of our work activity and interaction with others, there is an opportunity to take advantage of integrating the diverse amount of context available to the computer. That is, since the computer is becoming a dominant tool in which people process information and accomplish work, logs of its usage can be used to construct a substantial amount of context for a person’s activity. As the computer also becomes integrated in mediating communication among people, it can share that

context among people establishing contact in ways that support the contact negotiation process.

A prototype that we developed in 1995 that explores some of this potential is the Piazza system [Isaacs et al. 1996]. Piazza provided some awareness among people who were “working nearby” virtually. That is, they were using their computers to work on related tasks, such as browsing the same web page or editing closely related files of programming code. A strict definition of working nearby is to be using the same application to work on the same data at the same time, but relaxing those dimensions to include using similar applications to work on related data within some time range is even more interesting. Piazza offered awareness of people who were virtually working nearby, and made it easy to initiate communication (video conferencing, text chatting) among them. Since the computer was being used both to help people accomplish work and initiate contact, it could be used to detect and share a likely context for making contact.

In this way, the Piazza concept supports one aspect of greetings by sharing awareness of information that could suggest topic initiation. By “meeting” someone else working nearby in a computer workspace, one could anticipate that any communication initiated by that person at that time might be related to the common work context. We built a working prototype of Piazza, but only had limited experience using it. Our use confirmed its potential, but only began to explore how to define working nearby to provide enough useful encounters without also encountering too many distractions. The amount of internet connectivity that has emerged since then suggests an even larger opportunity of discovering meaningful opportunities of working nearby that could lead to making contact, which is ripe for further research.

More recently, we have been exploring a notion of *preset* and *pending calls*. This idea is driven by the observation that many contexts that anticipate initiating contact can be detected in the user’s activity on the computer. For example, if you receive an e-mail from someone, you might want to respond to it by calling her back on the telephone. Thus, we imagine a preset call where a speed dial soft button on your phone is programmed to the phone number of the person who sent you the e-mail. Preset calls help hasten contact initiation.

Pending calls build on a similar notion where the likelihood of a phone call is more strongly indicated. For example, research on IM usage [Nardi et al. 2000] suggests that an IM conversation often converts to a phone call once the IM has established the person’s conversational availability. A design implication that builds on that observation is to have the computer set up a pending call among people engaged in an IM where a

call is essentially placed to the other IM participants and put on hold. Thus, the IM participants could simply pick up the phone receiver to be instantly engaged in a phone call, without having to go through the mechanics of dialing a number, waiting for the ring, and ultimately having the other party pick up. Again, by recognizing the likelihood of initiating contact, the computing platform can be used to hasten the contact initiation process.

Our group's use of Awarenex over time has also led us to notice rhythmic patterns in people's activity that could be useful in coordinating contact. For example, since we on the West Coast could see when our East Coast collaborators were active on their computers day after day, we could begin to develop a model of when our East Coast colleagues typically break for and return from lunch, leave at the end of their day, etc. While this rhythmic awareness is usually shared among co-located workers, it could only be noticed among remote collaborators by using an awareness system over time. Begole et al. [2002] describe our studies that have explored these rhythmic patterns. Using the computing platform to integrate the logs of awareness data (when people are active on their computer) and the Awarenex Contact List interface that suggests when people are typically available for contact initiation presents a new opportunity for supporting contact initiation.

In this paper, I have shown how an understanding of human communication practices shaped the design of three communication prototypes. While each prototype provided new ways to support contact negotiation, there is still room for further research. While that the most recent prototype, Awarenex, in some ways provides the most support for contact negotiation, it does so without requiring the high fidelity of video connections. Instead, it uses text and icons to convey some of the cognitive and social cues for contact negotiation, without attempting to re-create the sensory cues through video.

Besides summarizing what we learned from studying these specific designs, I hope to have demonstrated the value of an overall design approach of building on an understanding of human practice. I have also suggested new directions for applying this design approach to develop new interfaces for mediating communication. Creating new interfaces based on an understanding of human activity allows us to design novel devices that enable people to do new things in familiar ways. As we invent new forms of computer-mediated communication, we need to focus on supporting contact negotiation to enable people to gracefully use these communication media.

## ACKNOWLEDGEMENTS

I am grateful to all the people at Sun Microsystems, Inc. who were involved with me in developing the prototypes described in this paper. David Gedye, Amy Pearl, and the Digital Integrated Media Environment team worked on the DCP. Rick Levenson, Monica Rua, Ellen Isaacs and the Conferencing and Collaboration team helped develop Montage. Nicole Yankelovich and Bo Begole were part of the Network Communities team that developed Awarenex user interface and provided valuable comments on earlier drafts of this paper.

## References

- Bannon, Liam (Ed.), *Journal of Computer-Supported Cooperative Work*, 3:1, 1995, pp. 29-96.
- Begole, James "Bo", John C. Tang, Randall B. Smith, and Nicole Yankelovich, "Work rhythms: Analyzing visualizations of awareness histories of distributed groups", *Proc. of the Conference on Computer-Supported Cooperative Work (CSCW) 2002* (New Orleans, November 2002), ACM, pp. 334-343.
- Bly, S., A., Steve R. Harrison and S. Irwin, "Media Spaces: Bringing People Together in a Video, Audio, and Computing Environment." *Communications of the ACM*, 36(1), pp. 28-47.
- Buxton, William A.S., "Living in Augmented Reality: Ubiquitous Media and Reactive Environments", in *Video Mediated Communication*, K. Finn, A. Sellen, & S. Wilber (Eds.) Hillsdale, N.J.: Erlbaum, 1997, pp. 363-384.
- Cassell, J. and H. Vilhjálmsón, "Fully Embodied Conversational Avatars: Making Communicative Behaviors Autonomous." *Autonomous Agents and Multi-Agent Systems* 2(1), 1999, pp. 45-64.
- Clark, Herbert H., "Chapter 18: Language Use and Language Users", *The Handbook of Social Psychology*, G. Lindzey and Elliot A. (Eds.), New York: Harper and Row, 1985, pp. 179-229.
- Dourish, Paul and Sara Bly, "Portholes: Supporting Awareness in a Distributed Work Group", *Proceedings of Conference on Computer-Human Interaction (CHI) '92*, (Monterey, CA, May 1992), ACM, pp.541-547.
- Erickson, Tom and Wendy Kellogg, "Social Translucence: An Approach to Designing Systems that Mesh with Social Processes," *ACM Transactions on Computer-Human Interaction*, 7(1), 2000, pp. 59-83.
- Fish, Robert S., "Comparison of Remote and Standard Collaborations," *Proc. of the Conference on Technology and Cooperative Work*, Tucson, Arizona, February 1988.
- Fish, Robert S., Robert E. Kraut, Robert W. Root, Ronald E. Rice, "Evaluating Video as a Technology for Informal Communication", *Proc. of the Conference on Computer Human Interaction (CHI) '92*, (Monterey, CA, May 1992), ACM, pp. 37-48.
- Flores, Fernando, Michael Graves, Brad Hartfield, and Terry Winograd, "Computer Systems and the Design of Organizational Interaction", *ACM Transactions on Office Information Systems*, Vol. 6, No. 2, April 1988, pp. 153-172.
- Goffman, Erving, *Behavior in Public Places: Notes on the Social Organization of Gatherings*, Glencoe: The Free Press, 1963.
- Heritage, John, "Conversation Analysis", in *Garfinkel and Ethnomethodology*, Cambridge: Polity Press, 1984, pp. 233-292.
- Isaacs, Ellen, John C. Tang, and Trevor Morris, "Piazza: A desktop environment supporting impromptu and planned interactions", *Proc. of the Conference on Computer-Supported Cooperative Work (CSCW) 1996*, (Boston, November 1996), ACM, pp. 315-324.
- Isaacs, Ellen, Alan Walendowski, and Dipti Ranganathan, "Hubbub: A sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions", *Proc. of the Conference Computer-Human Interaction (CHI) 2002*, (Minneapolis, April 2002), ACM, pp. 179-186.
- Kendon, Adam, "Behavioural Foundations for the Process of Frame Attunement in Face-to-Face Interaction", in *Discovery Strategies in the Psychology of Action*, G.P. Ginsburg, M. Brenner, and M.V. Cranach (Eds.) London: Academic Press, 1985, pp. 229-253.
- Kendon, A. and A. Ferber, A Description of Some Human Greetings. in *Comparative Ecology and Behaviour of Primates*, R. P. Michael and J. H. Crook (Eds.) London & New York: Academic Press, 1973, pp. 591-668.
- Kiesler, Sara, Jane Siegel, and Timothy W. McGuire, "Social Psychological Aspects of Computer-Mediated Communication", *American Psychologist*, 29(10), October 1984, pp. 1123-1134.
- Knapp, Mark L., Roderick P. Hart, Gustav W. Friedrich, and Gary M. Shulman, "The Rhetoric of Goodbye: Verbal and Nonverbal Correlates of Human Leave-taking", *Speech Monographs*, 40, August 1973, pp. 182-198.
- Kraut, Robert E., Carmen Egido, and Jolene Galegher, "Patterns of contact and Communication in Scientific Research Collaboration", in *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*, J. Galegher, R. E. Kraut, C. Egido (Eds.), Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 1990, pp. 149-171.

Nardi, Bonnie, Steve Whittaker, Erin Bradner, "Interaction and Outeraction: A Study of Instant Messaging in the Workplace", *Proc. of the Conference on Computer-Supported Cooperative Work (CSCW) 2000*, (Philadelphia, December 2000), ACM, pp. 79-88.

Nelson, Les, Sara Bly, and Tomas Sokoler, "Quiet Calls: Talking Silently on Mobile Phones", *Proc. of the Conference on Computer Human Interaction (CHI) 2001*, (Seattle, April 2001), ACM, pp. 174 - 181.

Parkes, A. M., "Drivers business decision making ability whilst using carphone", in *Contemporary Ergonomics, Proc. of the Ergonomic Society Annual Conference*, 1991, London: Taylor & Francis, pp. 427-432.

Pedersen, Elin Rønby, "Calls.Calm: Enabling Caller and Callee to Collaborate" *Proc. of Conference on Computer Human Interaction (CHI) 2001, Extended Abstracts*, (Seattle, April 2001), ACM, pp. 235-236.

Sacks, H., E. Schegloff, and G. Jefferson, "A simplest systematics for the organization of turn-taking for conversation", *Language*, 50, 1974, pp. 696-735.

Schegloff, Emanuel A. "Sequencing in Conversational Openings", in *Directions in Sociolinguistics: The Ethnography of Communication*, J. J. Gumperz & D. H. Hymes (Eds.), New York: Holt, Rinehart, & Winston, 1972, pp. 346-380.

Schegloff, Emanuel A. and Harvey Sacks "Opening up Closings", *Semiotica*, 1973, pp. 289-327.

Siegel, Jane, Vitaly Dubrovsky, Sara Kiesler, and Timothy W. McGuire, "Group Processes in Computer-Mediated Communication", *Organizational Behavior and Human Decision Processes*, 37, 1986, pp 157-187.

Suchman, Lucy, "Do Categories Have Politics? The language/action perspective reconsidered", *Journal of Computer-Supported Cooperative Work*, 2, 1994, pp. 181-190.

Tang, John C. and Ellen A. Isaacs, "Why Do Users Like Video? Studies of Multimedia-Supported Collaboration", *Computer Supported Cooperative Work: An International Journal*, 1(3), 1993, pp. 163-196.

Tang, John C., Ellen A. Isaacs, and Monica Rua, "Supporting Distributed Groups with a Montage of Lightweight Interactions", *Proc. of the Conference on Computer-Supported Cooperative Work (CSCW) '94*, (Chapel Hill, NC, October 1994), ACM, pp. 23-34.

Tang, John C. and Monica Rua, "Montage: Providing Teleproximity for Distributed Groups", *Proc. of the Conference on Computer Human Interaction (CHI) '94*, (Boston, April 1994), ACM, pp. 37-43.

Tang, John C., Nicole Yankelovich, James "Bo" Begole, Max Van Kleek, Francis Li, and Janak Bhalodia, ConNexus: Extending awareness to mobile users, *Proc. of the Conference on Computer Human Interaction (CHI) 2001*, (Seattle, April 2001), ACM, pp. 221-228.

Whittaker, S. "Theories and Methods in Mediated Communication", in *The Handbook of Discourse Processes*, A. Graesser (Ed.) Cambridge, MA: MIT Press, in press.

Whittaker, Steve, David Frohlich, and Owen Daly-Jones, "Informal workplace communication: What is it like and how might we support it?", *Proc. of the Conference on Computer Human Interaction (CHI) '94*, (Boston, April 1994), ACM, pp. 131-137.