Fugue:
A Computer Mediated Conversational System that Supports Turn Negotiation

Tara Rosenberger Shankar, Max VanKleek, Antonio Vicente, Brian K. Smith
MIT Media Laboratory
20 Ames Street
Cambridge, MA 02139
Phone 617-253-4564
Fax 617-258-6264
{tara, vankleek, int18, bsmith}@media.mit.edu

Abstract

Computer mediated communication (CMC) systems are providing new ways to communicate. Yet many text-based CMC systems do not represent the invisible, interactive practices, such as turn-negotiation, commonly found in face-to-face (FTF) conversations. Designing new text structures may help address these problems. To explore the effects of adding rhythmic, non-verbal cues to computer-mediated communication (CMC), we developed Fugue, a networked environment that creates visualizations of conversations as they occur, dynamically highlighting social presence and turn-negotiating events on a two-dimensional grid. In this paper we examine traditional text-based conversation systems, the role of turn-negotiating, social presence and activity in FTF communication, and how Fugue makes these non-verbal linguistic cues explicit within a graphical display.

1. Introduction

Hollan and Stornetta caution computer mediated communication (CMC) designers to focus upon communication needs fundamental to human interaction [13]. Highly ranked amongst their list of communication needs are negotiating *turns,* how speakers take turns at talk*,* and *repair*, how speakers correct for speaking and communication mistakes. Turn negotiation and repair are invisible, interactive FTF practices, and not easily represented in a visual medium such as text.

We believe that in order to perform turn taking and repair, conversationalists require a high degree of conversational feedback about others' temporally based behavior. People contextualize their talk primarily through prosody, the "musical" attributes of speech – melody, dynamics, rhythm, tempo and pause [7]. As a channel, the telephone affords speakers prosodic cues, including rhythm, whereas text CMC rests solely upon syntactic structure. We are interested in designing a text CMC system that represents *rhythmic cues*. We expect this to enable speakers to perform turn negotiation and repair. Lacking these cues, people are left with only the linguistic units to help them understand what another person intends. This is often insufficient.

The current design of text-based CMC systems does not lend itself to rhythmic representation. Text-based CMC systems often use the text layout developed for early manuscripts. Alphabetic characters usually flow from a left column border to a right column border, forming a line of text. These lines are then stacked from top to bottom of a page. When computer user interfaces came into being, the designers did not reconsider text design from the perspective of what a programmable interface can afford. They adopted the ancient left to right, stacked text design. This ancient design cannot carry enough rhythmic information. We need to explore temporal typographic solutions for real-time text CMC.

2. What is Rhythm?

In face to face (FTF) conversation, people respond to one another in a continuous manner. They are able to perform synchronously through their understanding of contextual cues, one of which is rhythm. In FTF conversation, rhythm is enacted paralinguistically and kinesthetically [10]. Rhythmic paralinguistic cues include when one begins to talk, where one pauses and for how long, how one paces different parts of an utterance and gives certain units greater emphasis [4,6]. Kinesthetic cues include when and how one nods, gestures, and shifts gaze [1].

People use rhythm within speech to contextualize their words. For example, a person will increase their pace of speech while lowering their intonation pitch to signal their intention to finish their turn [16]. Furthermore, rhythmic cues are usually not redundant

with respect to linguistic semantics [7, 11]; the same words said with different rhythm and stress are interpreted differently.

Since we are working with text input to a CMC system, we focus only on the timing of how people pace the words within an utterance, and that utterance relative to others' type-written talk. How they would intone or stress the utterance is unknown.

2.1. Turn Taking or Negotiating?

We distinguish between turn-taking and turn-negotiating. The difference between taking and negotiating involves being aware of others' presence and state of action. This awareness is important in the design of text systems.

Turn-taking does not require participants to coordinate their interaction through an awareness of what others are doing. IRC and email exchanges do not broadcast how speakers develop their turns; therefore, no other conversationalist can guess at how the "speaker" might complete the turn, choose to interrupt them or tightly chain on their utterances! Without witnessing the development of a turn, there can be no turn negotiation. As a result, interruption is impossible to perform in turn-taking environments.

Turn-negotiating, then, is having an awareness of others' current state of action. This awareness provides the context in which a group of people can decide who gets the role of speaker at any given point in time. There are implicit guidelines for how people negotiate turns speaking [20]. The guidelines are based upon a knowledge of syntax, semantics, and local contextualization cues. As such, conversational groups enact polite turn negotiating in ways other groups would find rude [10]. Conversationalists also learn to project how and when a speaker is going to complete her|his turn. Listeners can interrupt on the basis of this projected knowledge.

In order to create a text CMC environment in which turn negotiating can occur, conversationalists must first be able to see a speaker develop an utterance. In other words, a CMC system that wants to support turn-negotiation must first support a high degree of temporal feedback.

Speakers can usually expect to yield their turn at talk at the first possible ending projected by a listener, a *transition relevance place* [14]. The closer a speaker gets to a potential ending, the easier it is for listeners to project what the speaker is going to say, and begin speaking [7]. In figure 1, "B" and "R" both use overlapping turn structures because of their ability to project the other's intention.[1]

Overlapping Speech

R: so you understand =the requirements=
B: =yeah, I under=stand them
R: so you understand the requirements?
B: =yeah, I understand them
R: =and the schedule?
B: yeah

Figure 1. The words enclosed within the equals sign (eg.=) were spoken by the two participants simultaneously [10, p.247].

Overlapping talk occurs frequently within FTF conversation. It is used as a cue to politeness, and also to signal the emotional engagement of the conversationalists. If someone begins speaking at a juncture in someone else's speech that can not be interpreted as a transition relevance place, the overlap is considered an interruption; if the speaker overlaps at a transition relevance place, it is considered enthusiastic [21]. One's use of rhythmic coordination affects others' perception of one's character. A CMC system that supports turn negotiation must be able to support overlapping talk.

Backchannel signals are feedback cues that hearers give speakers [27]. "Mmmmm", "ok", "yes", nodding or other body gestures are common backchannel cues. They inform the speaker that the listeners are indeed listening, how the listener is interpreting the talk, and even to encourage the speaker to continue speaking rather than giving up their turn at a transition relevance place. Backchannels often overlap other talk, but rather than competing for the floor, cede it. Backchannels are a fundamental cue to conversational continuity and flow.

In figure 2, Cedric encourages Martin to continue through a relevant transition by saying "mmm," usually an affirmative backchannel cue. If Cedric had paused for an extended moment prior to giving the "mmm" backchannel to Martin, rather than overlapping with his speech at the transition place, then he would communicate something like hesitancy of support or belief. The rhythm by which people

[1] The form used in figure 1 is not well suited for seeing both how multi-party conversation overlaps and proceeds continuously over time. A number of computerized notation systems have emerged in the past ten years for documenting the simultaneous modal channels (speech, prosody, gesture, face, body) across multiple people (for example, [17]). These systems are both inspirations for Fugue's design, and a mirror of Fugue's horizontal, scrolling interface.

perform turn negotiation is important to the interpretation of the person's intentions.

Backchannel Speech

Cedric: =Mm
Martin: =But in the meantime it'd cost Keegan three spots for the feature.
Cedric: Yeah?
Martin: So, boy when Keegan come in he, you know how he's got a temper anyway,
he just sc=reamed his damn engine you know,
Cedric: =Mm

Figure 2. Martin holds the floor in this conversational excerpt. Cedric's brief "Yeah?" and "mm's" are examples of backchannel speech [22, pg. 216].

Repair in FTF conversation is the ability to correct for speaking mistakes in pronunciation, syntax, or propriety. Repair in FTF conversation is done publicly; both the mistake and the repaired version are available to all conversationalists (although most pronunciation repairs are not even consciously processed by listeners). The chat model presumes that pronunciation repair is synonymous to editing. Therefore, all chat writing can be checked off-line before it is sent to the public forum. If a CMC system is to support turn-negotiation, it must allow repairs – effectively typing edits – to be made within the public forum, so as not to interrupt the temporal flow of conversation.

2.2. Presence, Activity and Attention

Some of the most fundamental communication needs in a conversation are cues of *social presence*, *activity* and *attention*. Cues to *social presence* [23] can be seeing people around you or hearing someone breathing on the phone line. For example, when someone you are talking with on the telephone stops talking for more than two seconds, you say "hello?" You use their reaction to make judgements about their *social presence*, *social activity*, and *attentional* state. Respectively, are they still there? How involved are they in this activity? And, what or who were they paying attention to?

Every CMC system makes (consciously or otherwise) design choices about how to represent social presence, activity and attention. Chat represents social presence in one of two ways. One way is intermittent: when one's message is visible in the scrolling text window, one "exists" in the conversation. Or, more recent chats have separate windows in which are listed the names/aliases of everyone currently logged into the conversation. Chat gives no concise visual indication to how active those people are and have been over time (recent systems described in [8, 25] that incorporate graphics are notable exceptions).

3. Text Conversation Systems

Text CMC systems have different ways of broadcasting and representing text. We describe the design of UNIX Talk and chats as distinct archetypes in text CMC design.[2]

UNIX's Talk interface is divided into two halves, and each person's utterances are displayed within one half of that screen region on a keystroke by keystroke basis. The conversational feedback of how a speaker paces their talk is as high as text CMC systems allow. In other words, a computer system with keyboard input cannot send a smaller linguistic unit than a single alphabetic character. By sending each keystroke as it occurs, Talk conveys the highest possible rhythmic cue in text-based CMC systems.

UNIX Talk stops short of representing how speakers' turns interrelate temporally. It constructs a conversation by treating each person's composite utterances as a wrapped text body, and then stacking those texts on a single screen. Users must construct the turn-taking patterns by noticing when the other person begins or stops typing in their text buffer. As well, the representation of history maintains no visual clues to the turn structure that occurred. This perceptual constraint puts a severe limitation upon the number of people pragmatically possible in a Talk conversation.

Ytalk (figure 3), a more recent version of Talk, theoretically allows unlimited numbers of people to converse. ICQ (figure 4) is an Internet version of Ytalk with a "windowed" arrangement of people's talk.

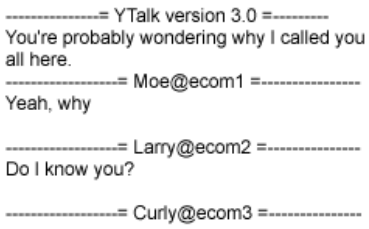

Figure 3a. Example of four participants beginning a conversation within YTalk 3.0. Participants have an individual buffer in which they can type.

[2] We refer to both IRC and Java-based Chats as "chat" systems, since their interface designs are very similar.

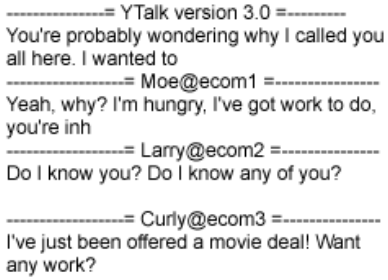

Figure 3b. As talk moves to the second line, it becomes visually confusing to notice when someone begins and finishes talking, as well as the temporal synchronization that would indicate who is talking in response to whom.

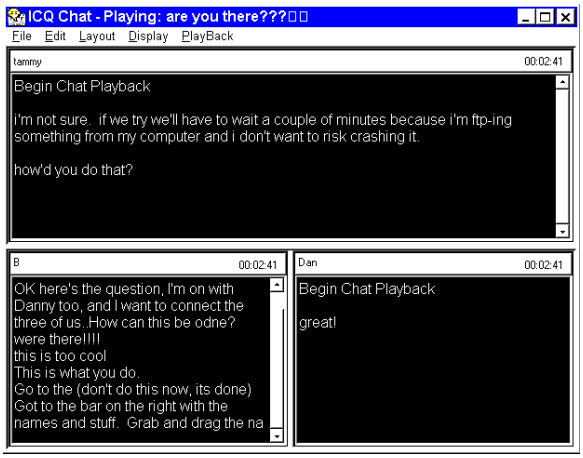

Figure 4. ICQ Chat gives each participant a window in which their words are placed, keystroke by keystroke, like UNIX Talk.

In chat systems, the utterance is not displayed until the speaker hits a carriage return. Other conversationalists never see the development of the speaker's turn, only the product of the speaker's turn. As such, Chats afford a lower level of conversational feedback than UNIX's Talk function.

Each participant's utterance is added to the bottom of the screen. As the number of participants increases, it becomes more difficult to understand the cohesion between the turns [19]. In fact, multiple conversations may actually be conducted simultaneously within the same IRC conversation. A reply in a busy Chat may appear many turns after the last applicable message (see Figure 4), making the written output appear cacophonous. Conversely, a Chat interface can actually accommodate tens of people simultaneously, something UNIX Talk could not accomplish. However, average turn lengths become shorter as more people join [5], suggesting a possible parallel drop in the ability to express a complex thought.

Figure 5. An IRC conversation formats a conversation like a scrolling document. The system places each person's turn at the bottom of the screen. A Java Chat uses the same username plus phrase format as IRC does. It differs from IRC by including a list of participant names and "prettier" networking prints.

4. Representing Conversation

In an effort to improve upon the limitations of Talk and Chat we discuss, many researchers and commercial enterprises have developed innovative UI structures for conversation. A recent approach has been to represent conversational settings and people with virtual avatars. In the most behaviorally realistic avatar systems, avatars can use hand and eye gestures at rhythmically appropriate moments, as humans might [26]. Yet, these spatial environments often use a Chat-like interface to communicate, thus incurring similar problems of not being able to negotiate turns.

Some text-based CMC systems provide conversational structure by forcing speakers to label their words as belonging to a particular class of speech act [8, 2]. Similar approaches have channeled conversation through consecutive tasks [17] or required conversation to revolve around specific subjects [19] and threads [24]. These systems provide subject cohesion, but at the cost of limiting how a conversation may develop. Because their feedback rate is approximate to a Chat model, turn negotiation may not occur within such environments.

Recent systems have abstracted the FTF context, either making the spatial distances between the visual representations of people a functional element of propagating the words they say [25], or making the textual interaction functionally control an abstract representation of the conversational group [8]. Both systems focus on representing spatial conversational cues. But their use of Chat networking structures does not permit them to support turn negotiation.

Little work within CMC has focused upon supporting the *interactive and invisible practices* that people implicitly perform in FTF situations, such as

turn negotiating. To make these invisible, temporally based practices visible, we need models for representing multiple voices as continuously as possible through time.

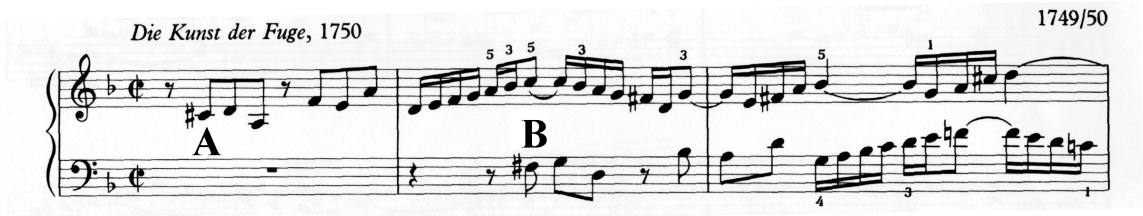

Figure 6. A single staff system of a score is shown. The staff system is an excerpt of a Bach fugue [3]. The fugue motif is introduced at point A; at point B the same motif begins a fifth step lower.

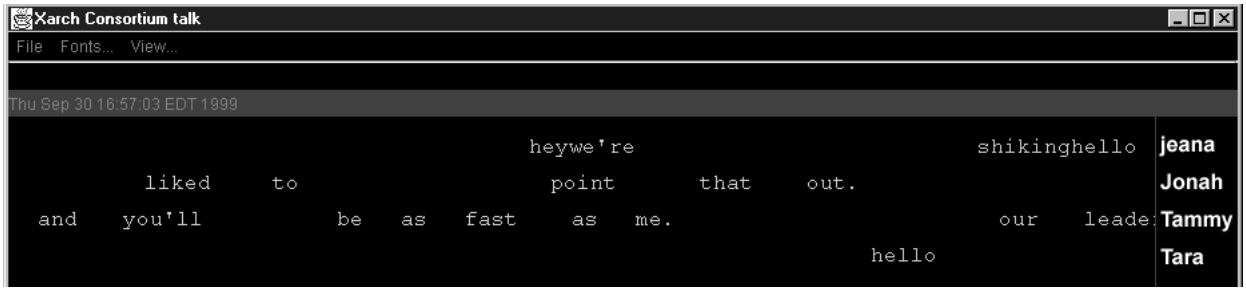

Figure 7. This is an excerpt from a four-person conversation occurring in Fugue between jeana, Jonah, Tammy and Tara. The participants are greeting and joking with each other, and commenting upon the novelty of the interface. These overlapping turns could not occur within a Chat model of conversation.

5. Fugue: Conversational Music

Fugue makes features of FTF conversation like social presence, activity and attention explicit within an interface while structurally permitting simultaneous talk. We believe that by giving people the knowledge to negotiate conversation, through a contextual awareness of what others are doing at a moment to moment basis, more rich and complex meanings can be socially created and exchanged in text CMC. Different turn structures are a part of the product of complex conversational meaning making.

The inspiration for Fugue comes from musical notation, where multiple instruments are represented in a single space. Inscribed upon musical staffs in a left to right manner are tonal events and their relative rhythms (see figure 5). If the events occur in the same vertical space, they are to be played simultaneously. If the events follow one another in the horizontal space, they are to be played consecutively. Fugue is a real-time recording of a conversational performance. In Fugue, the temporal relationships exhibited as speakers communicate – such as simultaneous speech found in interruptions and backchannels – are preserved with vertical overlap, like musical notes on a score. Conversation in which turns chain forward are noted in Fugue along the horizontal continuum. Therefore, in this way Fugue creates a "conversational score" of a conversation.

A musical score is composed of numerous staff systems stacked onto a page. Fugue differs from the musical score by not breaking the temporal (horizontal) dimension into pieces and stacking them like a text. Rather, Fugue allows the horizontal dimension to continue for as long as the conversation. In a musical score, you can read music that is noted earlier in the piece by turning back the pages. To read conversational history in Fugue, one scrolls back through the conversational score.

5.1. Social Presence in Fugue

In music notation, each space and line that comprises the musical staff correspond to a unique note value. Similarly, each Fugue participant has a horizontal graphic space within the Fugue "staff." Any words that appear within that space can be assumed to be that person's words. These horizontal spaces lie one on top of the other. When a participant first logs into an existing conversation, a new text line shows up underneath those that were already there; the staff becomes one space bigger. Her or his name is printed towards the right-hand side of the text space as a permanent background feature. This line of text is the mark of a participant's social presence within a

Fugue conversation. It never disappears so long as they are logged into the conversation.

Fugue represents social presence and activity using text design and layout. When someone logs in to Fugue, they are given a new horizontal space in which the words they type will be placed. It is labeled with their name. Similar to text practices, Fugue places this new textual line, representative of a new conversationalist, underneath the existing participant lines. This horizontal space serves to broadcast their social presence, similar to a person's entry to a room. Like an avatar, as long as a person remains logged into Fugue, their allotted space never disappears, even when a participant has not talked for hours.

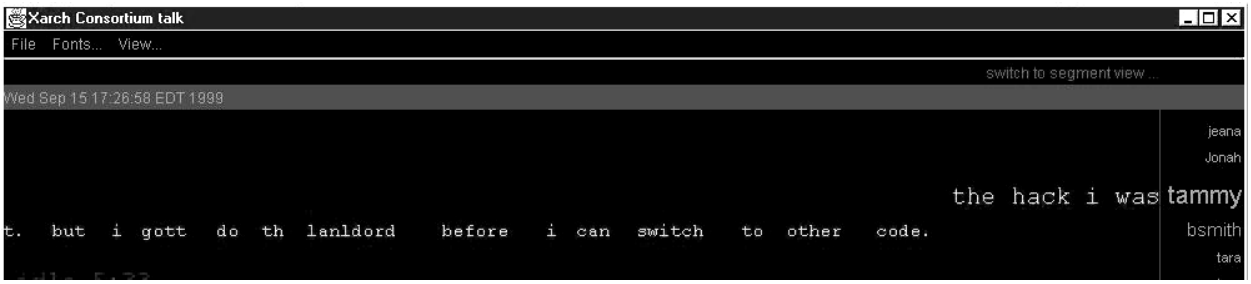

Figure 8. In this conversation, Tammy holds the floor. Since she is an active speaker, her text is maximum size. bsmith has spoken most recently and as such, his text is second largest.

5.2. Activity Cues in Fugue

To represent a speaker's changing level of activity, Fugue keeps track of the amount of time elapsed since each participant's last keystroke. This amount is used to determine the current vertical size of a participant's line of text. Three visual states are possible. When the participant is actively typing, the text is full-size. When they have just discontinued typing, the text slowly minimizes over time. When they have been silent for a while, the text is minimized to a point where it is just readable. As soon as a participant types a key, their line immediately becomes full size. Fugue never removes a line. In this way, representations of social presence and activity level are de-coupled. Each person is visible even if they say nothing; yet by talking their text becomes more animated. In figure 8, Tammy's text is largest because she currently holds the floor. Bsmith's text is smaller because of his most recent status as speaker; the others' lines are minimized to reflect their status as listeners.

Activity level is also apparent by the rate of text flow. Fugue broadcasts each keystroke of each participant. Each character typed is broadcast and displayed within each Fugue client's horizontal text line as close to real-time as the network allows. The entire staff of text flows like a ticker tape from the right to the left, as the conversation grows. The rate of text flow usually depends upon the speed of the fastest typist currently typing, not the sheer passage of time. When people type simultaneously, the text may flow even faster. In other words, a conversation does not flow unless at least one person is saying something.

5.3. Directing Attention

Changing the visual prominence of text focuses others' attention. While a participant is talking, their text remains large. The text lines representing people who have become listeners slowly minimize, making the speaker's text the largest. When a participant who was silent speaks, their line gets big, attracting the gaze of other participants.

If a participant's inactivity level reaches some system defined critical amount, the amount of elapsed time since they last spoke is printed in their conversational line. Others can use this number to make a judgment about the level of attention the participant is likely giving the conversation. For example, if someone left their computer to eat dinner but remained logged into a Fugue conversation, the amount of time since their last keystroke (e.g. "IDLE 0:5:01") appears in a dull color in their conversational line.

Figure 9. Fugue conversational window involves two main sections: the active conversation that can be read and moves along the horizontal, and the "zoomed out" view, or the iconic view**, of the overall conversation. The iconic view acts as a navigation device to read the conversational history, as well as informs one about the kind of turn-taking patterns that have occurred.**

6. Conversational History

Fugue's UI portrays the history of a conversation in two ways: (1) a window that shows a length of speaker turns within the conversation, usually the most recent, and (2) a global view of the entire conversation that shows turn lengths and levels of participation across time.

In global view, Fugue continues to represent each participant as a (potential) horizontal space, and a thin line drawn within that horizontal space represents the presence of an utterance (as opposed to silence). We intend the global view to have two primary functions. The first allows one to judge how many people are present, who has been talking the most, who has not talked, and how long a participant's turns are relative to other participants. Each time there is a change in the participant context (e.g. someone logs in or out of the conversation), a new conversational segment boundary is drawn. The iconic view's second function is to allow a user to navigate back through conversational history. A square box pictured in the global view indicates the portion of conversation currently visible. One can scroll through the conversational continuum using the global view, changing the readable portion of conversation in the active view.

6.1. Building History Asynchronously

Fugue enhances its FTF model of turn-negotiating with an asynchronous messaging component. Here we describe how we implemented support for both, and why conversation needs both asynchronous and synchronous technologies.

Fugue speakers can direct an email message to a specific point in the temporal history. By clicking on the conversational grid at a point, a message window opens in which one may type an extended message. After sending it, the message appears in the conversational grid as a round message marker (see figure below). By rolling a mouse over this message marker, the conversational grid fades into the background, and the entire textual message appears as an overlay.

There are public and private types of messages. If you click on someone else's grid line, that person alone will receive the message. If you click on the line above the entire conversation, it is a public message that anyone participating in the conversation can see. In figure 9a and 9b below, Tara sends a message to the group recommending further action on naming a developing technology.

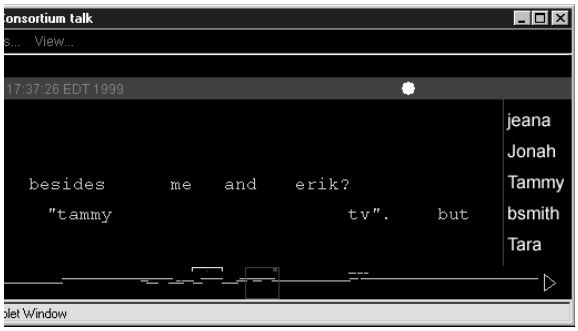

Figure 10a. The round message marker in the grey bar above the active conversation indicates that someone has added a public note to history. If it were a private note addressed to you, the round marker would have appeared within your line.

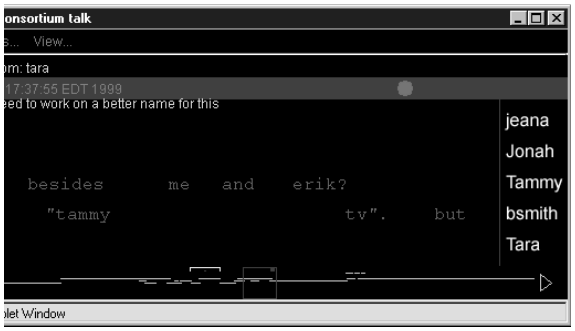

Figure 9b. When the mouse rolls over the message marker, public or private, the message body appears in the window.

Asynchronous messaging allows conversationalists to make conversational *repairs* on a conceptual level, as well as comment in a thoughtful "off-line" manner after a real-time conversation has occurred. The difference between these two actions is that a repair is a comment upon one's own words, and an off-line comment is a comment upon someone else's words.

Conversation is often based upon referencing, and then re-interpreting, repairing, or expanding, previously uttered words. FTF conversation is ephemeral; hence this is an action based only upon one's ability to recall the words and context. In Chat systems, one can scroll back through the transcripts in order to refresh one's own memory of what transpired, even to quote exact words. In Fugue, one has the opportunity to make fresh comments that reference an exact moment in history by leaving a message. This message marker, indicating "read" or "unread" status, is also visible in the global view.

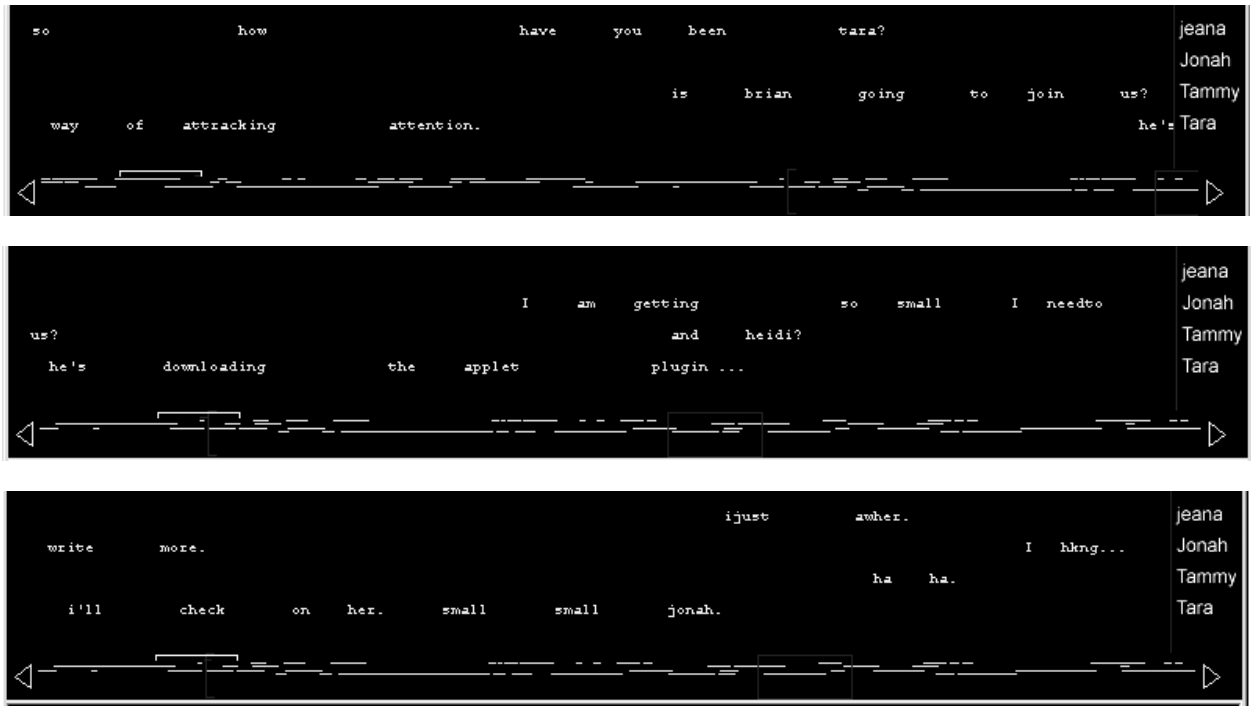

Figure 11a, b, c. This conversation excerpt is captured from the beginning of the group conversation. It is a representative sample of the type of chorus-like interaction that occurred before the more formal meeting began. There is a high level of turn interruption and overlap during this opening portion of the conversation.

7. Evaluation Discussion

Fugue gives people the potential to talk simultaneously, and then makes visible the turn-negotiating patterns they choose to enact. There is both evidence for and against Fugue's structured display of turns developing in turn influencing how people develop their turns. We do see different kinds of turn negotiating patterns. Interruption is common, similar to FTF conversation. But people also tend to complete their turns after being interrupted, differing from FTF conversations. Turn lengths vary dramatically across different types of interactions. Power relationships that one would find in FTF conversations do carry over into the Fugue environment. We discuss these results below.

We generated conversational data by conducting a group meeting within Fugue, involving five participants, for an hour. The conversation consisted

of two types of interactions. The first was organizational and informal, when people were gathering together and preparing for the meeting to begin. The second was the heart of the meeting when the interaction was more formal, and generally only one person talked at a time. The results we discuss below compare and contrast both interactions.

The conversation before the meeting began was chorus-like with short turns that overlapped one to the next continuously. One participant talked of the collaborative, chorus-like interaction pictured in figure 11, "There's a vitality to it, like being on a river…it was like a race. When people learned to coordinate, it was fun." Presumably, "learning to coordinate" meant chaining talk so closely – if not overlapping – that the talk never stopped flowing. There seemed to be a uncomfortable feeling when the talk stopped, almost akin to an uncomfortable period of silence between topics in a spoken conversation.

Fugue Meeting Conversation Excerpt

Tammy: is there enough time to get that going and get it out in the real [1]world before the october meeting?

Brian: [1]It'll be tight, but it's a necessary evil. Else I'm not sure [2]what the point of showing it would be.

Jonah: [2]What age kids do you want to use?

Brian: I have a group of high schoolers. They're standing by….for months now….

Figure 12. Numbers in brackets indicate when an overlapping turn begins. Members of the group discuss putting together a user test of experimental software.

The conversation excerpt in figure 12 demonstrates turns that chain forward through overlapping beginning and endings occur, but with less overlap relative to the length of the turn than the conversation in figure 11. There are long stretches when only one person is talking. People seem to listen to one another, even during long stretches of talk. Evidence for people attending to one another's long stretches of talk can also be found in the unanticipated use of the cue to speaker activity. To cue the speaker that they were listening, listeners would occasionally press the space key to gain speaker activity status for a moment. Their name and talk line would gain maximal size, and then shrink down again. They did not type anything in addition to the space key; so the speaker could not interpret their action as requiring a response, or desiring the floor.

Interruption in Fugue does not earn the same response that interruption in FTF conversation does. People do not give up the floor. They finish their turn after being interrupted, even when knowing that the interrupter perfectly grasps what they intended to say. If figure 12 were a transcript of speech rather than Fugue talk, Tammy would most likely have stopped talking after "real world…". Yet everyone who is interrupted continues their thought to some syntactic conclusion, plus final punctuation. This may indicate that in text CMC, people adjust their interaction to both the people, and to the written history being developed.

People were frustrated at the inability to interrupt someone else's talk *successfully*. In other words, people could interrupt – and did – but their effort did not gain them the floor. Rather, as in the figure 12, the interrupted person would complete her/his thought even though another had interrupted. In one hour of talk, there is only one instance when someone leaves their thought unfinished – and that is likely because it was the professor who interrupted them. Power relations found in FTF conversation can also be found in the structure of turns in Fugue.

Students would talk simultaneously for up to fifteen words. This was often where divergent topics of conversation were simultaneously proposed. Rather than the simultaneous talk itself being a battle for the floor as in FTF conversation, the floor was determined after the speakers concluded when the listeners chose to pursue only one.

Although the outcome of interruption does not follow FTF practices, people valued knowing they had been interrupted. A few participants complained of not noticing when someone interrupted them because they were busy typing on – and looking at – the keyboard. (Fugue is easier for the touch-typist, who can watch the screen while talking.) Non-touch-typists were disturbed to know that they had not been aware of an interruption, and thus had not responded socially. This suggests that the awareness of being interrupted does affect their conversational behavior in some way, even though it appears that people continue their thought with no regard for the interruption.

8. Conclusion

We have presented an overview of Fugue, a networked environment that supports conversation between multiple participants. Fugue supports a high level of rhythmic conversational feedback that allows people to negotiate turns. It creates a context through explicit, visual representations of overlapping "speech", interruptions, and back-channel cues that participants can read in a manner similar to a musical score. These rhythmic coordination cues are built synchronously during conversation, preserved graphically in conversational history, and added to

asynchronously. Fugue transcripts exhibit interactions that can be identified through unique turn structures. Interruptions in Fugue occur commonly, but lead to behavior that differs from interruptions in FTF conversation.

9. References

[1] Auer, P., di Luzio, A. (eds.). 1992. *The Contextualization of Language*. Amsterdam and Philadelphia: Benjamins.

[2] Auramaki, E. et. al. 1988. A Speech-Act-Based Office Modeling Approach. *ACM Transactions on Office Information Systems*, Vol.6, No.2, pp126-152.

[3] Bach, J. S. 1975. *The Fugue*. Ed. and Annotator Charles Rosen. London, Oxford University Press.

[4] Beckman, M. 1986. *Stress and Non-Stress Accent*. Dordrecht, Holland/Riverton: Foris Publications.

[5] Cherny, L. 1999. *Conversation and Community*. Stanford, CA, CSLI.

[6] Couper-Kuhlen, E. 1993. *English Speech Rhythm: Form and Function in Everyday Verbal Interaction*. Philadelphia, Benjamins.

[7] Couper-Kuhlen, E., Selting, M. (eds.) 1996. *Prosody in Conversation: Interactional studies*. New York, Cambridge University.

[8] Erickson, T. Smith, D. N., Kellogg, W. A., Laff, M. R., Richards, J. T., and Bradner, E. 1999. Socially Translucent Systems: Social Proxies, Persistent Conversation, and the Design of 'Babble.' In *SIGCHI*, ACM Press.

[9] Flores, F. et. al. 1988. Computer Systems and the Design of Organizational Interaction. *ACM Transactions on Office Information Systems*, Vol. 6, No. 2, pp. 153-172.

[10] Gumperz, J.J. 1982. *Discourse Strategies*. Cambridge: Cambridge University.

[11] Gumperz, J.J. 1992. Contextualization and understanding. *Rethinking Context.* Eds. Goodwin and Duranti. Cambridge, Cambridge University.

[12] Gumperz, J.J., Berenz, N. 1993. Transcribing Conversational Exchanges. *Talking Data: Transcription and Coding in Discourse Research.* Eds. Jane Edwards and Martin Lampert. New Jersey, Lawrence Erlbaum Associates, Inc., pp. 91-121.

[13] Hollan, J., Stonetta, S. 1992. Beyond Being There. *The ACM Conference on Computer Human Interaction*, May 3-7, pg. 119-125.

[14] Jefferson, G. 1973. A case of precision timing in ordinary conversation: overlapped tag-positioned address terms in closing sequences, *Semiotica* 9, 47-96.

[15] Kendon, A. 1990. *Conducting Interaction: Patterns of behavior in focused encounters*. Cambridge, University of Cambridge Press.

[16] Lehiste, I. 1980. Phonetic characteristics of discourse. Paper presented at the Meeting of the Committee on Speech Research, Acoustical Society of Japan.

[17] McDaniel, S., Olson, G., Magee, J. 1996. Identifying and Analyzing Multiple Threads in Computer-Mediated and Face-To-Face Conversations. *ACM Conference on Computer Supported Collaborative Work*, Cambridge, MA.

[18] McIlvenny, P. 1990. Communicative Action and Computers: Re-embodying Conversation Analysis? *Computers and Conversation*. Eds. Luff, Gilbert, Frohlich. New York, Academic Press.

[19] Roddy, B., Epelman-Wang, H. 1998. Interface Issues in Text Based Chat Rooms. *SIGCHI Bulletin*. Vol. 30, Num. 2, pp.119-123.

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

[21] Schegloff, E. 1987. Analyzing Single Episodes of Interaction. *Social Psychology Quarterly*. 50(2), pp.101-114.

[22] Schegloff, E. 1992. To Searle on Conversation: A note in return. *(on) Searle on Conversation*. Eds. Searle et. al. Philadelphia, Benjamins.

[23] Short, J., Williams, E., Christie, B. 1976. *The Social Psychology of Telecommunications*. New York, John Wiley & Sons.

[24] Terveen, L., Hill, W., Amento, B., McDonald, D., Creter, J. 1997. Building Task-Specific Interfaces to High Volume Conversational Data. ACM *SIGCHI*, 22-27 March.

[25] Viegas, F., Donath, J. 1999. Chat Circles. In Human Factors in Computing Systems: *The Proceedings of CHI '99*. ACM Press.

[26] Vilhjálmsson, H., Cassell, J. 1998. BodyChat: Autonomous Communicative Behaviors in Avatars. *Proceedings of ACM Autonomous Agents '98*, Minneapolis, May 9-13, pp. 269-276.

[27] Yngve, V. H. 1970. On getting a word in edgewise. Papers from the Sixth Regional Meeting of the Chicago Linguistic Society, April 16-18. Chicago, University of Chicago, Department of Linguistics. pp. 567-578.