Medium Effects: Turn-Taking and Back Channels

Text-based CMC does not offer the same channels available in face-to-face conversation. Yet it is apparent that text-based communication is both useful and productive in the MUD community, despite the poverty of the text channel. How does text conversation differ from face-to-face conversation? How do users have to adapt to make it comfortable and to obviate misunderstandings? How does the mode of communication affect the register at an interactional level?

The opportunity for smooth turn-taking, interruption, requests for clarification, and speaker change have been identified as important aspects of interactivity (Whittaker 1994) in CMC and more generally for communication of information (Clark & Brennan 1991). Feedback signals of understanding and attention, which have been called back channels (Yngve 1970), have also been deemed important in conversation (Clark & Schaefer 1989). This chapter explores the effect of the MUD medium on turn-taking and back channel behavior during conversation, suggesting that users as rational actors modify their behavior according to the demands of the medium.

5.1 MUD Media Characteristics Revisited

As I discussed in chapter 2, media systems differ in how they support interactivity in Rafaeli's (1988) sense: the "expression of the extent that in a given series of communication exchanges, any third (or later) transmission (or message) is related to the degree to which previous exchanges referred to even earlier transmissions" (p. 111). I related the idea of interactivity to Goffman's notion of "reciprocally sustained involvement" (1957, p. 50), and the notion of "recipient design" in conversation analysis (e.g., Schegloff 1977). Although much work on media system characteristics has come from a technological determinism perspective, I stressed the fact that describing media characteristics involves acknowledging the
social context of use. Towards a typology of system descriptors, I suggested that we consider message granularity, both from the system's architecture perspective and the user culture perspective; whether message storage occurs and for how long; the temporal structure of the transmissions (asynchrony, synchronous, two-way); and the channels available, both for human-human communication and human-computer communication. Systems support greater or lesser degrees of telepresence, the feeling of being in another location, depending on their interface characteristics.

Special turn-taking behaviors, adapted to media characteristics, are a part of mediated communication registers (Gibbon 1985, Smith 1979). In CB, turn-exchange signals have evolved to facilitate use of the simplex (one-way) channel where no interruptions are possible and only one may speak at a time (Smith 1979). In the MUD register, the "loses" routine (see chapter 4) evolved from the fact that any number of people may be typing at once, but messages are only transmitted upon pressing return. The message granularity and lack of two-wayness in the channel were the source for this routine, in which players communicate that their typed utterances are redundant because another player has uttered a similar thing first.

The MOO media characteristics are summarized here:

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Breakdowns</th>
</tr>
</thead>
<tbody>
<tr>
<td>message granularity</td>
<td>technically: any length line may be sent</td>
</tr>
<tr>
<td></td>
<td>socially: between 5 and 13 words</td>
</tr>
<tr>
<td>message storage</td>
<td>archived: MOO mail</td>
</tr>
<tr>
<td></td>
<td>archived (only in local buffer): real-time conversation</td>
</tr>
<tr>
<td>temporal structure</td>
<td>synchronous, not two-way</td>
</tr>
<tr>
<td>channels</td>
<td>visual (text only)</td>
</tr>
</tbody>
</table>

MUDs may be characterized as having generally small grained messages, usually from a word to a clause in length, which are ordinarily not stored by the system (so impermanent), but users themselves may keep logs, or maintain large local "scrollback" buffers of conversation, as discussed in chapter 3 (which effectively act as temporary storage facilities). The user transmits conversation a line at a time by pressing return. Message granularity in the synchronous conversation mode is actually largely up to users, who learn to respond dynamically to the constraints of the context, as I discuss in this chapter. Long messages decrease the sense of co-presence and awareness of others in the medium, by decreasing the real-time feel (since composing a long message takes more time). MUDs are synchronous but not two-way channeled, so interruptions cannot occur in the sense of overlap.

Asynchronous modes of communication in the form of mailing lists and private mail exist in some MUDs as well, supporting larger grained messages which are stored. Many of the political actions on LM, for instance, occur via mailing lists. The synchronous mode of
communication in MUDs will be my focus in this chapter, however.

5.2 Previous Literature

According to the Computer Supported Cooperative Work (CSCW) literature and literature on CMC, one major problem posed by mediated interactions is the difficulty with turn exchange. Tools that support group collaboration at a distance, "groupware," are particularly prone to problems caused by the mediated situation. Synchronous video and audio connections have not been as beneficial as expected (Egido 1988, Heath and Luff 1991a). Visual gestures are not perceived as readily in video environments, and ability to discriminate among speakers decreases as group size increases (Watabe et al. 1990). Furthermore, the sorts of floor behavior found in face-to-face conversation don't appear to be well-supported by voice channel (e.g., asides, conversants pursuing sub-conversations; McKinlay et al. 1994).

Floor control is a particularly contentious topic in the groupware literature. Some authors advocate strict one-at-a-time speaker/actor roles (Novick and Walpole 1990), while others condemn this structure as too restrictive (Ellis et al. 1991, Beaudouin-Lafon and Karsenty 1992). Woodburn et al. (1991) found that strict sequencing in dyadic text-based CMC situations increased the amount of time for task completion, and users generally preferred to determine their own strategies for turn-taking. In the full-duplex split screen connection they studied, users liked being able to see messages being composed character-by-character, and could frequently begin answering in parallel.

Control flow depends heavily on properties of the medium. McKinlay et al. (1994) note that properties of CMC may offset some of the problems that increased group size causes in face-to-face situations; because of the permanence of the medium, users depart from serial turn-taking and adopt topics in parallel (cf. Black et al. 1983, McCarthy et al. 1993). Face-to-face models of turn-taking may not be appropriate for CMC situations, they conclude.

Black et al. (1983) studied turn-taking in email discussions versus in classroom conversations. They also found that the time-delay in email resulted in multiple topics (or "multi-threads") being pursued in parallel, instead of one-at-a-time as in face-to-face conversation. Real-time messaging, in their study, displayed single-topic threads, like face-to-face conversation. They also found back channeling to be minimal in email, resulting in “dangling” conversations with no closure.

McCarthy et al. (1993) report on text-based conferencing in which users can only see contributions after a typist presses a "send" key. Since composition was not visible and upcoming messages were not signalled, users missed contributions or wasted time typing what another user had already sent. Users were also unsure when another user was "listening" and available for conversation. Timely repair mechanisms were also missing, leading to persisting misunderstandings. McCarthy et al. conclude that message structuring to help
with signals of understanding (e.g., back channel response buttons), to establish referents, and to enforce response to queries might be helpful additions to text-based CMC systems.

Oviatt and Cohen (1991a, b) found that keyboard communication of instructions to assemble a water pump resulted in fewer back channels, longer turns, and fewer clarifications or interruptions. The medium is deemed less efficient than the interactive speech medium, since the task also took longer to complete. However, the task involved was a very hands-on task, so the keyboard modality was not optimal in any case. Their finding is also predicted by Clark and Brennan (1991), who attach high cost to reception, production, and speaker change in the keyboard modality.

In the sections below, I challenge some of these conclusions, in particular the claims that fewer back channels necessarily result from keyboard interactions, that longer turns result, and that repair is impeded by the medium. I also discuss the effect of the permanence of the medium on the topic threading, and show that multi-threading is still common, despite the real-time nature of the conversation in MUDs. Contrary to the finding in McCarthy et al. (1993), users in MUDs have adapted to the problems caused by the one-way channel in which messages are not received until a user finishes typing and sends them. Sending smaller grained messages and using frequent back channels are some of the ways in which they have adapted.

5.3 Turn-Taking in the MUD

The entire notion of turn-taking as a process in conversational interactions stems from an "economy" model of conversation, in which both interlocutors want access to the floor and must negotiate for it. Turn-taking as conceived in Sacks, Schegloff, & Jefferson (1974) is an aspect of synchronous communication, in which all potential speakers have access to the channel at once, and for processing reasons, only one speaks at a time, ideally. It is viewed as a competitive process.

Process aside, the turn itself is a matter of some theoretical and empirical concern. Goodwin (1981) discusses examples like this (p. 18):

Example (1):

1 John: Well I, I took this course
2 (0.5)
3 Ann: In how to quit?
   [ ]
4 John: which I really recommend.

Silence within a turn is usually called a "pause," while silence between turns is a "gap." The silence in line 2 above is classified differently by the different speakers: for Ann, it is a pause, for John, it is a gap. John's continuation in line 4 reveals that he was not finished
with his turn, but to Ann, he appeared to be. Her line 3 therefore becomes an interruption, since John is not finished.

Is the turn an intentional unit? John’s intention was for lines 1 and 4 to constitute a single turn. However, for Jaffe and Feldstein (1970), what occurs is the turn structure; they simplify the procedure for determining “possession of the floor” according to which speaker is speaking. Simultaneous speech they resolve by fiat: a “speaker switching rule” is invoked to determine who wins the competition. Although such a process may be useful descriptively, it is apparent that turn-taking and turn-identification are complex processes that are problematic for speakers as well as discourse analysts. As Goodwin (1981) suggests, the turn itself is a matter of empirical and analytic concern, rather than just a predefined category useful in analysis of other structures.

Murray (1989), in a discussion of a computer messaging system and turn-taking, considers the turn an intentional unit. Turns may be made up of utterances and messages; a message is the line of text sent by the message system (which has an upper limit on size of message), and an utterance consists of the messages sent by a speaker between messages from other speakers or the system. In this example, P and T are different speakers:

Example (2):

T1: IF YOU HAVE A NAIVE USER WITH A HOME TERMINAL,
    YOU REALLY HAVE A PROBLEM
P1: It’s not just home terminal old, it’s the whole general
    question of orderly communications/information for users.
T2: I THOUGHT THE OLD STUFF WAS PRETTY WELL DOCUMENTED.
P2: ... batch going away is a case in point, we’ve got to provide info
    -- rather detailed info ahead of time.
P3: yes, and i want to put out similar documentation for home terms.
P4: also batch.

Messages P1 and P2 are a turn, to Murray (pp. 324-325), because P2 is an elaboration of P1, and does not refer to T2’s content. P3 and P4 are part of a new turn. P2, P3, and P4 are all one utterance, because they were sent without interruption. Turn allocation does not occur as in face-to-face speech; parties can take turns at will.

In the discussion that follows, I will use the term “utterance” to refer to Murray’s messages: what a user sends in a single command line (using either emote or “say” or other communicative commands) to the MOO server. Each utterance appears only after a user is done typing, presses return, and the network (which may be subject to the delay the community calls “netlag”) gets it to the server and to the other connected users. There is no constraint on message length in MOO, unlike in Murray’s message system; users determine their message (or utterance) length based on the demands of the conversational context, as I will discuss below.

Because size of utterance is entirely determined by a user, I believe it to be difficult to separate out an intentional turn unit that is different from the speaker order that appears
on the screen (which is in part determined by the server and network timing, and thus out of a speaker's hands anyway). "Interruptions" clearly do not have the status they have in face-to-face speech, since overlap is not possible. If a speaker chooses to express a point in several small utterances, another speaker may take turns in between, at any point, and they are not as potentially damaging to the flow of conversation as in face-to-face speech. An analyst might choose to say that each utterance ends with a turn completion point or a "transition relevance point," but since any other speaker may speak at any time, with her utterances showing up between other speakers' utterances as the network gets them out, transition relevance points do not seem an appropriate notion for the medium. (I will expand on this topic in the next section.) Rather than talk in terms of "turns" as intentional units, I believe the topic-driven notion of holding the "floor" is more relevant here, given that multiple speakers may all type at once in the MUD medium (Edelsky 1993, Shultz et al. 1982, Hayashi 1991). I will outline this view after discussing the process of turn-taking described in Sacks et al. (1974) and contrasting MUD conventions with it.

For the sake of discussion below, I will consider any stretch of utterances by a single typist before another typist's utterance(s) to be a turn. So in the example below, lines 1-2 are a turn, lines 3-4 are a turn, and so forth.

Example (3):

1 Ted runs around in circles
2 Ted says, "I get to meet mi-chooEEeooelee"
3 Karen hehs.
4 Karen says, "you could've two days ago, but nooo00000000000000"
5 Ted shakes Karen.
6 Karen grins.

5.3.1 Sacks et al. (1974) Contrasted with MUD Turn-Taking

In this section, I compare claims about face-to-face turn-taking with that observed in the MUD. My benchmark for face-to-face conversation is the Sacks et al. (1974) article on turn-taking.

In face-to-face conversation, Sacks et al. observe, speaker change recurs. Overwhelmingly, one party talks at a time. Occurrences of more than one speaker at a time are common in face-to-face conversation, but brief. On the MUD, speaker change occurs if both parties are active at once. Sometimes people leave messages for idle people in their buffers, which they will see later, taking advantage of the persistent aspect of the medium. Interaction in real time is the norm, but it is not necessary. It is trivially true that only one speaker speaks at a time in the MUD, since only one person has access to the channel at once. No text can overlap, although speakers may be typing at the same time. The order of utterances can be mixed up, however, if someone is responding to a remark that has since been superseded by another remark, and the response appears to follow the wrong utterance.
Unlike in face-to-face conversation, multiple threading can occur in conversations in the MUD, as Black et al. (1983) describe occurring in email conversations. MUD conversation is more like the real-time messaging system that Black et al. looked at and found lacking in multiple threading, however. Since in face-to-face encounters speech participants are usually only involved in one encounter at a time, Black et al. conclude that time delay causes the multiple-topic threading. Since MUD conversation shows some multiple threading, I conclude that it is the persistence of text on the screen that makes multiple threading possible (as well as the lack of interruptions due to the one-way, half-duplex mode), rather than the time delay. McCarthy et al. (1993) and McKinlay et al. (1994) support this claim. MUD conversation does not generally show multi-topic utterances as Black et al. (1983) observe in email messages. Each utterance is primarily one topic, directed to one person.

Example (4):

1 The Swedish guest says, "how do I get out?"
2 Tom [to Shelley]: so woj said something about dinner
3 lynn says, "quit"
4 Shelley [to Tom]: cool, i was hopgin he'd find you
5 The Swedish guest says, "thanks"
6 Mike says, "happy birthday!"
7 Shelley is still waiting for her orders
8 Shelley grins.
9 Shelley [to Mike]: thanks
10 Mike | The guest [to Swedish guest]: ok
11 Tom [to lynn]: so the big deal about amulet, if it delivers, is portability
12 Mike [to Tom]: you put him up to that, I bet
13 The guest leaves the library to the east.
14 lynn nods to Tom.
15 Tom says, "well, that combined with a high clue quotient on the part of the designers"
16 Tom laughs at Mike.

In this example, several conversations are occurring at once. I answer the guest’s question about how to leave in line 3, she thanks me in line 5; Tom and Shelley discuss her birthday dinner plans; and Mike, referring to a private joke with Tom about directed speech, quotes from an earlier line in the conversation with a “paste” (vertical bar) and then refers back to the line with a deictic pronoun in his next utterance, line 12. Meanwhile, another conversation between Tom and lynn starts up about a graphic user interface toolkit called Amulet. Such multi-threading is reminiscent of collaborative floors in Edelsky (1993), as I will discuss below.

In face-to-face conversation, transitions from one speaker to the next with no gap or overlap are common. A slight gap or overlap or neither characterizes the majority of transitions. (Yngve 1970 appears to conclude otherwise, although based on much smaller sample
size.) In a MUD, timing is unreliable, so close timing only occurs if both parties are active and responding immediately to each other's comments. There is usually a small gap between an utterance and the response to it; however, due to the need to type the whole utterance before sending it to the server.

In face-to-face conversation, turn order is not fixed, but varies. Turn size also varies. Turn size and order vary in the MUD too. Interestingly, short turn size often makes conversation more interactive. An example of short turn size follows:

Example (5):

1 Tom says, "ROBOT KIDS"
2 lynn says, "?"
3 Tom says, "hm, you didn’t hear about that one?"
4 lynn dun think so
5 Tom unfortunately doesn’t remember it very well; Shelley or Jon might well have a log.
6 Tom says, "i’m sitting around in my room on lambdamoo one day"
7 Tom says, "and some guy teleports in"
8 Tom says, "because i’m a wizard"
9 Tom says, "and introduces me to his rl fiance"
10 Tom says, "and i kind of watch helplessly"
11 Tom says, "as the situation gets weirder and weirder"
12 Tom says, "and at one point he’s asking me"
13 lynn nods.
14 Tom says, "or telling me"
15 Tom says, "that he and his fiance would like to use lambdamoo to create robot kids"
16 Tom says, "it’s that weird, and he’s not self-conscious about it at all"

Tom appears to break his turns at what would be prosodic boundaries in oral speech, or what Sacks et al. might consider transition relevance points. However, as mentioned above, this notion does not map well onto MUD conversation, since any speaker may begin typing an utterance during any other speaker’s turns, and no overlap occurs. My back channel nod in line 13 could have occurred anywhere in between his numerous turns. The same speaker, however, breaks his turns at very different places in the next example. The context for the discourse below is a discussion of the origins of EM, which is supposedly inspired in part by interactive fiction ("the i.f. stuff"); like Zork or other text-based adventure games. A new character had just entered the room, and some discussion of his unusual name, Michele_Amercia, followed.
Example (8):

1 lynn [to Tom]: what do you work on when you work on the i.f. stuff?
2 Tom says, "i was talking about the history of EM and stuff, M_A"
3 Michele_America says, "oh... yes, you're right but this is my name. :)"
4 Tom [to Michele_America]: well, there's this long-time regular user here named Michele, so your name kind of gives us a jolt because of its obvious similarity.
5 Penfold says, "and 'america' is kind of an unusual last name in the USA"
6 Michele_America says, "yes... my name is Michele because i'm italian. but i live in Portugal."
7 Tom [to lynn]: ha. building and its infrastructure—integration, english stuff. and occasionally we talk about agents or weather or simulations or something, which i think would also contribute.
8 Michele_America says, "and the America is because my great-grandfather was an emigrant in the USA. :)"
9 Tom says, "but as you know, we generally devote almost no time to actually building stuff, so that leads to a reputation for being against or uninterested in it, and so on and so on and so on"

It seems likely that Tom breaks his utterances in this example according to who he is speaking to, and what points he is making, rather than at clause or phrase boundaries as in the previous example. This probably makes the interleaved conversation easier to follow. Notice that in line 7 he uses directed speech (aka “stage talk”) to establish who he is speaking to, since gaze is not available. (In some MUDs without this capability, a speaker needs to address an individual by name, generally.) He does not redirect his remark in line 9, and we assume he is speaking to the same person as in his previous line 7.

In face-to-face conversation, length of conversation and topic are not specified in advance. This is true in the MUD as well. Furthermore, there are communicative options other than “say” and “emote” in the MUD, allowing distinctions and input sources not possible in
face-to-face conversation. In example (7), Tom pastes (the vertical bar indicates text is being pasted) from another source in order to illustrate what he is trying to program.

Example (7):

Tom says, "suppose i have oh here"
Tom | (define-macro (define-binding object event . program)
Tom | ' (set! (cmds object) (cons (lambda ()
Tom wants that to be (lambda () thestuffin' program)"

The “thought bubble” is another frequently used textual device in the MOO: “Karen o 0 ( once ).” The Antisocial commands, or interactions with objects, all involve communicative behavior that is further mediated by the MUD server, producing “utterances” not typed directly by the player (see chapter 4).

In face-to-face conversation, the relative distribution of turns is not specified in advance, it is an interacational achievement. This is true in MUDs as well. The number of parties in a conversation can vary in both face-to-face conversation and in MUDs. In the MOO I observe, there is some evidence that for more than 6 speakers talking within one minute, the number of words per minute that can be produced and understood drops. The graph in Figure 5.1 shows dense, representative, one-minute time slices from conversations with different numbers of participants, analyzed according to median number of words in the minute, number of utterances per minute, and words per utterance per minute. The number of utterances per minute climbs as more speakers participate (see graph in upper right corner), while the number of words per utterance decreases (see lower left). The intersection of the two tendencies results in a high point at 6 people, for words produced per minute (see upper left). The final box in the lower right shows the number of cases used for each data point; note that above 7 people, it becomes very difficult to find examples of that many people talking during one minute. There are certainly cases of 7 or more people being active in conversation at once, but usually not all are speaking frequently enough that they show up within a one minute time slice. For conversations like those, however, it becomes difficult to identify active participants, since many people may only speak occasionally.

Figure 5.2 shows a graph of turn size in a long conversation (98 minutes) between me and Tom. The histogram below it shows the distribution of different turn sizes. In Örestrom's (1983) study of the turns in the London-Lund corpus of spoken face-to-face conversation, he found that two-thirds of all turns were below 20 words long, and the median turn length was 13 words. We can see that well above two-thirds of the MUD turns were less than 20 words long. Figure 5.3 shows the turn length in both dyadic and group

---

¹My logs are time-stamped. In order to generate the data for this and other graphs, I used perl scripts to parse my log files.

²The single point for 10 speakers is anomalous in several ways: it is rare for 10 people to be active, and in this particular case, it was an interview situation in which a journalist was probing about the history of EM and its relationship to LambdaMOO. The conversation was therefore more focused than usual for large party discussions, with a single topic controller most of the time.
Figure 5.1: One Minute Time Slices During Conversations with Different Numbers of Conversants
conversations; I considered 6 conversations each of 6 (or more) people speaking contrasted with 2 people speaking; the number of turns in the dyadic conversations ranged between 32 and 285 turns, and the number in 6+ parties ranged between 169 and 458 turns. The graph shows the average of the medians for the turn lengths in each conversation type, including high and low median figures to illustrate the range for each type. The average (of the median turn lengths) for dyadic conversations is 7.8 words; for groups, 5.5 words. The comparatively large range of variation for dyadic conversation is probably partly due to how distracted some speakers were (resulting in shorter turns perhaps) or the topic of conversation (more task-oriented or focused conversation probably results in longer turns). In groups, individual differences in concentration on the topic are equalized. These results about turn size contradict the prediction in Clark and Brennan (1991) and the findings in Oviatt and Cohen (1988) that turn size should be longer in keyboard modalities than in voice. There may be a difference because the task-oriented exchange in Oviatt and Cohen (1988) resulted in long explanations, of a one-sided nature, rather than highly interactive spontaneous speech. My findings support the Rational Actor Hypothesis from chapter 2: users adapt their communicative behavior for the medium; smaller turn size, hence smaller grained messages, increases interactivity, the degree to which messages may be related to earlier messages. Shorter messages allow more opportunity for back channel and repair, as I discuss in the next section.

Talk can be continuous or discontinuous in face-to-face conversation, where interruptions may occur. This is also true in MUDs, but in different ways. People can be interrupted "in real life," but the interruption isn’t visible in the MUD the way it is in face-to-face speech. Conversants need to plan for this eventuality. Example (8) illustrates a person distracted during conversation. The connection messages give an indication of the times events occurred. In line 7, Daphne notices Mike is “at work” and therefore prone to interruption.

Example (8):

1 Mike says, "I might end up in Newark for grad school."
2 < connected: 'draw (#2533) on Mon at 14:50.
3 Daphne says, "Really? Which one?"
4 Daphne says, "I like that this moo starts out on a highway"
5 Daphne says, "Wow, you’re a wizard? Oh yes, you told me that but I’d forgotten. Can you make me a programmer?"
6 Daphne pokes Mike
7 Daphne says, "Um, you must be at work"
8 < disconnected: Honda (#82) on Mon at 14:55.
10 Mike blinks.
Figure 5.2: Plots of Turn Size in Dyadic Conversation with Histogram of Sizes
Figure 5.3: Median Turn Sizes in Group and Dyadic Conversations, 6 Conversations Each (with High and Low Medians Shown)
Mike says, "Hi, sorry, yeah, people came in to ask me a question."

Turn allocation techniques are obviously used in face-to-face speech. A current speaker may select another speaker, or self-selection may occur. Use of an adjacency pair is one way of selecting another speaker. The first pair part, e.g., a question or an invitation, requires a response from another speaker, an answer or an acceptance or rejection. Other examples of adjacency pairs include greeting/greeting, request/grant, offer/accept sequences. A speaker may use an address term (like a name) or gaze to select next speaker, along with the first pair part. Failure to respond to the first pair part is seen as the responsibility of the speaker selected for the turn.

In a MUD, adjacency pairs exist, but silence is not attributable to the respondent. As mentioned, an interlocutor may have been called away suddenly and be unable to answer. Use of the directed speech option functions like naming an addressee to help to pin down the intended respondent in an adjacency pair (contrast Murray 1989 who described messages sent to only one person and suggested that naming functioned in other ways when it occurred). It also functions like eye contact.

Example (9):

Ray [to Tom]: cafe
Tom [to Ray]: ok, can i demo "zork to lynn first?"
Ray says, "sure"
lynn says, "yay"

Confusion does occur when a speaker uses a second person pronoun and the intended referent is not clear from context, however. Note Shelley's use of directed speech to clarify her reference here.

Example (10):

Shelley says, "wait, you're not going now?"
Tom says, "who?"
Shelley [to Tom]: you, to the mfa place

Murray (1989) suggests that greeting and closing adjacency pairs don't occur in her message system data; in fact, they do occur in my MUD data (exemplified by Type 1 emotes, discussed in section 4.8). Murray hypothesizes that closings don't occur in her data either: because other media take precedence over the computer conversation, neither party has anything new to contribute to the topic; or the initial reason for opening conversation has been

---

*Yngve (1970) mentions that this is not a neat phenomenon, however; speakers may ask questions but retain the floor through subsequent utterances before turning it over to a respondent.
resolved. Phone interruptions or someone entering an office can prevent further transmission and lead to a lack of proper closings. While silence does result from interruptions to MUD conversations, it's not appropriate to say that the other media take precedence, precisely; generally someone interrupted "in real life" returns to the MUD conversation afterwards. Lack of new information or resolution of the purpose for the conversation as reasons for concluding conversation presuppose purely task-oriented use of the medium, which is not true of the community on the MOO I observe. People in EM are in a social context, occasionally having technical, task-oriented conversations there; the social framework is not forgotten when the task is concluded. Clearly, the motivation for use of a communications medium has a large role to play in the way it is used, and the results in Murray (1989) do not carry across to the MUD.

Other-initiated repair will select the previous speaker in face-to-face conversation. This may take the form of partial or full repetition of previous utterances, usually with question intonation, or questions about the content of a previous utterance. In example (11), from Sacks et al. (1974, p. 717), Lori asks a question that selects Ben as the next speaker; it is his responsibility to clarify his previous utterance.

Example (11):

Ben: They got—a garage sale.
Lori: Where.
Ben: On Third Avenue.

In a MUD, intonation isn't available, so questions must be explicit. Other-initiated repair does select the previous speaker. Here Hacker asks Ray clarifying questions, which he answers.

Example (12):

Ray builds the education offices
Hacker says, "This is such a good idea."
Hacker says, "Where are they Ray?"
Ray says, "by the pond"
Karen thinks it's a clumsy implementation tho
Karen [to Ray]: in moocity?
Hacker says, "Pond is w or e of LM?"
Ray says, "east"
Hacker says, "Oh, along the brown path then?"
Ray says, "yeah"
Hacker says, "Gotcha."

In face-to-face conversation, self-selection occurs during a pause, when the first starter gets the floor. (Note the relationship of this phenomenon to classroom situations, where the
first hand up gets to speak first, usually.) Sacks et al. (1974) point out that the need for planning utterances leads to lots of appositional beginnings like "well," "but," "and," "so..." Speakers use these as pre-starts (the way tags are post-completers) to get the floor. They then create an utterance that projects their turn length. Using question words indicates a question is coming, for instance; intonation can also signal the size of the upcoming turn. (Yngve 1970 apparently disagrees, however, about the importance of signaling turn size. He notes that interruptions do occur when turns are syntactically or intonationwise incomplete.)

There are lots of appositional beginnings in MUDs, despite the fact that speakers have more time to edit their responses and must plan their utterances in order to type them out. The appositional starts may be used to mimic face-to-face conversation processes, or to affiliate to previous turns. Brine, an LM regular who visits on EM occasionally, pointed out to me recently that "EMers" use "so" to begin a new topic quite often. In (13), Tom uses a free-standing appositional start at line 3, possibly to signal that more is coming.

Example (13):

1 Tom is thinking about putting a realtime parties tracker into loom.
2 lynn says, "what would that do?"
3 Tom says, "well"
4 Tom says, "initially, just maintain a window that shows you roughly what 'parties' does now"

As mentioned, "turn-constructional units" are employed in face-to-face conversation. Turns can be projected one word long, or phrasal, or sentential. Turns are usually attempted at "possible completion points" in sentences, clauses, phrases, and after some single words. Turn-completion points are normally syntactically sensible. In the examples below from Sacks et al., (p. 721) the overlaps occur at possible turn-completion points.

Example (14):

George: I saw em last night [at um school
Henry: [They're a riot.

A: Well we just wondered,
A: We just came in from Alexandria,
A: just got home
A: and [these winds were so bad we're gettin' scared again, hah
B: [Mm hm,
B: No, [we doh-
A: [And we wondered whether we should go to a motel or something.
B: No, you stay right where you are.

Back channels and repair attempts occur at possible turn completion points (or "transition relevance points") according to Schegloff (1982).
Speakers in the MUD have the option of "interrupting" by speaking at any point during another's talk; rather than overlap, their utterances will be interleaved with the other speaker's utterances. Back channel responses, or utterances that appear to function similarly, are often offered during another person's stretch of talk. Repair attempts for both content and disorderly talk also occur.

5.3.2 Repair Mechanisms

Repair mechanisms exist in both MUD discourse and face-to-face conversation for dealing with turn-taking errors and violations. In face-to-face conversation, some of the repair inventory consists of interruption markers ("excuse me"), restarts, repeats, and premature stopping. The repair mechanisms in the MUD differ slightly because of the lack of a two-way channel in the medium. The "loses" routine (discussed in chapter 4) is one common response to discovering that another speaker has already typed what one is typing; similar to premature stopping, the user types "loses" instead of finishing her utterance, and then sends the text to the server anyway.

Responses can get disorderly, and require "fixing" explicitly. In this example, Karen's back channel to "cool" appeared after the next remark from Damon, changing the entire meaning of her nod. Karen straightens it out by telling him what she was nodding to. (Note the dropped preposition, a very common convention.)

Example (15):

Karen [to Damon]: inspection cuz buying a house.
Damon [to Karen]: aha! cool!
Damon says, "wait, you're rich?"
Karen nods solemnly.
Damon says, "house + boston -> rich, right? ok, just checking"
Karen says, "uh"
Karen says, "no, i nodded cool."
Karen grins.
Is says, "houst + boston -> newly poor"

The ability to paste lines from other sources is used for re-ordering utterances as well. In the following example, Shelley pastes a line from earlier in the conversation that she didn't get to respond to, so that she can respond to it properly as innuendo. (Line 2 is a "spoof" by Mike; he is typing in a question from Sadness, which he expected her to ask. The appended "--Mike" alerts everyone that it is not actually her typing!)

Example (16):

1 lynn [to Mike]: we can do it later.
2 Sadness says, "can you go bowling with us on Thursday,"
Mike?" --Mike
3 Mike says, "Why sure, I'd love to."
4 Sadness says, "Oh gee."
5 lynn snorts.
6 Shelley | lynn [to Mike]: we can do it later.
7 Shelley woo
8 Sadness laughs.

In example (17), Tom uses a paste from earlier to explain what his pronoun "this" referred to.

Example (17):

Rob [to Tom]: if they don't know about quitting then they're likely best off using@quit
Tom [to Rob]: um, thanks.
Mike [to Robin]: what's new?
Tom isn't quite sure what he's supposed to conclude from this.
Mike says, "From what?"
Tom | Rob [to Tom]: if they don't know about quitting then they're likely best off using@quit

In sum, the turn-taking methods and methods for repair of disorderly turns in the MUD are different from those in face-to-face conversation, but clearly related. Conversation can be orderly or disorderly, and if it is disorderly, repair mechanisms come into play, which are specific to the medium, like signals of premature preemption, explicit explanations of disorderly lines, or pasting into focus in the conversation at a later point.

5.3.3 The Shared "Floor"

Models of turn-taking that apply to face-to-face conversation do not map well onto the MUD medium, however. Given that there is no competition for the channel per se, but rather competition for attention or control of the discourse, notions of shared or collaborative floor (Edelsky 1993, Shultz et al. 1982, Hayashi 1991) seem to be more helpful than the standard turn-taking literature. These notions also appear more useful for theorizing multi-threaded topic discourse, as well.

Edelsky (1993) relates dissatisfaction with the Sacks et al. (1974) turn-taking model even for studying face-to-face conversation. The model does not account for different speaker perceptions of "interruption" behavior: cases where multiple overlapping speakers are "on the same wavelength" for instance. Simultaneous talk in Edelsky's group data sometimes seemed collaborative and sometimes seemed more interruptive.
Edelsky criticizes the turn-taking literature for focusing on one-at-a-time speech in dyads, or in institutional settings like classrooms or therapy sessions. Phone conversation has been another source of data for conversation analytic studies of turn behavior, despite the (largely unspecified) effects of the medium on the interaction. Most of these studies do not acknowledge a speaker's own intuitions about having the turn, and certainly in the CA tradition, the analyst's intuitions are no more valid (Jefferson and Schenkein 1978).

Edelsky goes on to define turn as "an on-record speaking... behind which lies an intention to convey a message that is both referential and functional" (Edelsky, p. 207). Floor is "the acknowledged what's-going-on within a psychological time/space" (p. 209). Clearly these are difficult to determine from the observer's perspective, but they allow for the existence of non-floor-holding turns, a concept that appears to be needed for the analysis of back channels (see section 5.3). Edelsky then proposes the existence of both collaborative floors and single floors. Morgenthaler (1990) proposes as an intermediate type the interleaved floor in which individual speakers who are identifiable speakers jointly work out a topic (distinguishable from collaborative because single speakers are not identifiable in the "babble" of collaborative speech).

In looking for alternative characterizations of talk, Shultz, Florio, and Erickson (1982) expand on notions of participant structures (Philips 1972) and define four types of "floor" at family dinners. Hayashi (1991) builds on their work and Edelsky's, describing the floor as a community competence, which is cognitively developed while participants interact, providing a continually updated context for each subsequent interaction. Floor reflects social concerns like solidarity and power, cooperation and conflict, and the like, as well as more operational notions like topic change and support. Hayashi summarizes the following types of floor breakdown:

Example (18):

Single conversational floor
Single person floor
Prime-time at a time floor
Speaker-and-supporter floor
Less-active interaction
Active interaction
Non-propositional floor

Collaborative floor
Ensemble
Joint floor

Multiple conversational floor
Side floor and main floor
Main floor in parallel

At the broadest level, there is either a single main floor, or multiple floors occurring. Where there is one floor, that floor may be either constructed by one person in the main, or
by multiple parties. A single main speaker may hold the floor and be seldom interrupted, as in sermons or lectures, and thus be in a "prime time at a time floor." When listeners interact with the speaker, there may be speaker and supporter floors. In more active supporter floors, simultaneous talk and overlap occurs, as well as back channel responses; in less active supporter floors, back channels occur but little overlap or simultaneous speech occurs. "Non-propositional floors" occur when speakers entirely pre-occupied with their own thoughts create a self-centered floor that has nothing to do with listeners and the on-going floor, for a short time.

Collaborative floors result from multiple speakers contributing at once to the floor. Ensemble floors result from little simultaneous talk; the group achieves a rhythm with close timing, as they attend to a common goal and the topic at hand. Joint floors, on the other hand, show sudden topic changes, overlap and interruption, turn-failures. A multiple conversational floor occurs when multiple floors occur at once, and subgroups form. These may either occur in parallel (when speakers share multiple topic goals), or one may become a side floor. Side flooring may constitute an interruption to the main floor.

Considering these floor types in relation to the MUD medium, we can see that not all the distinctions can be appropriate in a medium where overlap is not possible. Subgroup formation is also not possible with physical proximity or volume adjustment, so if it occurs, it must be purely topic driven. Modifying the above diagram for what might be possible distinctions in a MUD, we see:

Example (19):

\[
\begin{array}{c}
\text{Single conversational floor} \\
\quad \text{Single person floor} \\
\quad \quad \text{Prime-time at a time floor} \\
\quad \quad \text{Speaker-and-supporter floor} \\
\quad \quad \text{Non-propositional floor} \\
\text{Collaborative floor} \\
\text{Multiple conversational floor} \\
\quad \text{Side floor and main floor} \\
\quad \text{Main floor in parallel}
\end{array}
\]

Interestingly, perhaps because there are few formal speech events in MUDs, I cannot identify examples of prime-time at a time floors in EM. Fanderclai (forthcoming) refers to attempts to lecture in the MUD medium and floor control attempts generally, "...the uselessness of delivering a lecture in what is designed to be an interactive environment, and the irony of programs that silence people in an environment that can let everyone talk and be heard at the same time." The medium is far from ideal for formal lectures or debates; without non-intrusive visual feedback, it is disconcerting to speak for any length of time to an audience that does not respond.

Speaker and supporter floors are common in dyadic situations, where listeners provide back channel responses regularly, and one speaker may be dominant. In this example, Tom is the main floor holder, with lynn, Ray, and Patrick being supporters.
Example (20):

Ray says, "you're talking about CSL the VR space?"
lynn says, "see paul's paper on video connectins and space etc."
Tom thinks things are going well, anyway, in general.
Tom [to Ray]: well, not "csl" per se.
lynn is relieved by that.
Patrick ah.
Tom says, "in fact, we're probably going to get rid of the virtual
   representations of each of the kind of corporate structures"  
Tom says, "create more anonymous spaces, which may eventually be
appropriated and sometimes named"

Ray
Tom says, "emphasis on providing a variety of types of space, and
finding out which ones get used"
Tom eyes Ray warily.
lynn nods, great.
Tom [to lynn]: i showed them your map.

Non-propositional floors are very common in MUD conversation, particularly because people are often prone to narrate their actions in real life as they occur (Type 4 emotes). Frequently these narrations or side-comments do not appear intended to start a new conversation topic. On the other hand, frequently they are intended to start a new topic, and they just aren't picked up by anyone. Who has the ability to start a new topic, who is likely to be responded to when they speak, has a lot to do with power and social influence in the community, I suspect. (I have heard some complaints from newcomers about being "ignored" when they ask questions or make comments.) In this example, Dorian is asking about the history of EM, and Ted is talking off-topic about something at his job (I do not mean to imply that Ted has no social influence; on the contrary, he is one of the more influential speakers as far as linguistic innovation goes).

Example (21):

1 Dorian [to Tom]: Cool stuff without programming? How?
2 Karen [to Ray]: didn't you call it a party?
3 Ray says, "well, yeah, that was my operational analogy"
4 Ted, o 0 ( cu -1 /dev/cua0 9600000 ? )
5 Ray says, "EM was a big really long house party"
6 Tom says, "jay drew an analogy to a party on several occasions"
7 Tom says, "this doesn't mean it was one, it meant it shared
   characteristics"
8 Dorian [to Tom]: Characteristics such as?
9 Ray says, "mind you, we had disagreements on what this meant and
whether it was valid"

10 Tom [to Dorian]: for example, in a perfect world, you would be able to make a dog and train it to do tricks without ever having to use eval or set properties.

11 Tom ""

12 Ted is not real excited about editing 198 records by hand out of userbase, nope.

13 Tom [to Dorian]: characteristics such as "if somebody’s obnoxious, you make them leave".

14 Ray says, "people you know show up, hang out at your place"

Collaborative floors are common. In the example below, the thread about Shelley becoming different appliances is jointly developed by all. Shelley is explaining why she copied verb code (the “start” verb) from her washing machine object onto herself, and thus has become the appliance (see Cheryn in press b for a discussion of automation and cyborg identities).

Example (22):

1 lynn [to Shelley]: so when we agreed to take over the appliance project, I am not sure this is what I thought would happen with it.

2 Shelley [to lynn]: well, this all started because I was working on the washer and didn’t want to lose the part of the start verb I had right

3 lynn nods

4 Shelley [to lynn]: then someone tried to start me and was disappointed

5 Tom [to lynn]: I was thinking that as she makes more appliances she should repeatedly take on the characteristics of her latest work.

6 Shelley [to lynn]: so I added the missing properties

7 Kelly nods, hopefully when you’re done you can stop being a washer and re-create the actual washer.

8 lynn says, "oh, cycle through the latest appliance."

9 lynn cools

Although the lines above are relatively long, it may be interesting to note that Edelsky (1993) found an average turn size of 6.5 words in collaborative floors; in Figure 5.3 I showed that group turn size is about 5.5 words in EM. Most group situations are collaborative or multiple floor situations.

Since subgroups aren’t easy to identify within a conversation (because there is no visual or spatial dimension), the differences between side floor and parallel main floors are harder
to identify. As a participant in the event, I would claim that the interaction between the
guest and the regulars about code in the middle of an ongoing interview conducted by
Dorian constitutes a side floor in the example below:

Example (23):

Karen [to Dorian]: what's your book going to be on?
Ray was thinking Lotus Notes, but either would do
Dorian [to Karen]: Everything in the MUOniverse, basically.
Karen [to Dorian]: big book
Ray spends way too much time thinking about Lotus Notes at work
Ray says, "it is Evil"
The corporate guest says, "do you guys see someone here that
may be interested in helping a guest with a code problem?"
Dorian grins Karen. "I'm working on narrowing it down. But
something like a history/ethnography/travelogue."
Karen [to the corporate guest]: ask away...
Ray [to the corporate guest]: hm, well, just ask

Side floors might also be constituted when whisper conversations start between two
participants in the main floor. It is fairly common for a secondary whisper conversation to
start and even to address the same topic as the main floor, although "in private."

A parallel main floor example is below. In the midst of a conversation about stalking,
a conversation about a file system is occurring between Pete and Molly.

Example (24):

Harry says, "What's a good way to get rid of some boring, very
cute girl who's stalking you?"
Harry says, "Kill."
Anne says, "the kiwi margaritas at that place are so yummy i might
just die."
Molly [to Harry]: Kiss her.
Lynn [to Harry]: ask her to leave you alone.
Pete [to Harry]: give her my phone number
Pete eyes himself warily.
Anne says, "ask ted."
Molly hehs.
Molly eyes herself warily.
The guest says, "hi robin"
Robin lol vange.
Anne says, "give her my phone number."
The guest smiles
Pete [to Molly]: what does df say about the disk?
Pete says, "owait, i spose it doesn't mount so you don't KNOW."
Robin [to Harry]: whyforyouwanttogetridofher?
anne [to Robin]: cause she's boring.
Molly [to Pete]: Not mounted eh.
anne says, "if she was cute and interesting, he'd have married her already."

In sum, various types of shared floors are possible, but the distinctions in face-to-face situations that Edelsky (1993) and Hayashi (1991) bring out are not all available, due to lack of overlap in the medium. Multiple participant floors are in fact easier to achieve than in face-to-face conversations, given this lack of overlap. The common multi-threaded conversation identified in Black et al. (1983) and noted in McKinlay et al. (1994) may be analyzed as examples of multiple floors occurring at once.

Despite this array of floor types and the earlier discussion of repair mechanisms in MUD conversation, I do not mean to imply that floor manipulation is simple or painless. I discussed some evidence that asynchronous modes of CMC may be prone to dominance by vocal minorities in chapter 2; synchronous modes of CMC may also suffer from the same phenomenon. A MUDder recently wandered onto an academic MUD and reported about a log she read of a group of teachers discussing ways to use the MUD in teaching: "It's been my most chillingly RL-like experience on a MOO so far, watching the meeting dynamics and seeing what was said. ... My overwhelming feeling was of being in a stuffy room in some vaguely uncomfortable sort of chair with a bunch of very stuffy people, the women having to get their comments in edgewise while a man felt it necessary to be a very hands-on moderator and two other men argued ubiquitously about their pet point long after the others seemed to want to move on."

Fanderclai (forthcoming) refers to automated control devices for floor control in classroom situations. There is an example of voluntary use of real-life meeting tactics in a MUD meeting in the log of a discussion on TinyHell II, from early 19904. A moderator called on characters who raised their hands (or paws, as the case was):

Example (25):

Stewy opts for giving Random the floor before answering fur. fur says, "like i said earlier, i can ask someone to go with mee to eat, so that makes it more social"
Random says, "You can do that anyway."
LongThorn raises his hand, now.
Moira raises her hand too.
Random takes advantage of having the floor to answer fur. :)
Stewy pulls out his notepad to record the queue of speakers.
Random says, "I should point out that if people (ie, The

---

4At http://www.ccs.neu.edu/home/lpb/muds/reality
Public) really had ANY interest in more realism in muds, they'd
bonnie well use the :command to simulate realism more often. It
is my observation that (with the possible exception of tinysex :) this is rare indeed."
fur raises his paws.
Random says, "I distinctly recall having a spaghetti dinner
on DaisyMud."

(The "." command is the emote command's shorthand.) MUD conversation can seem
chaotic, to new users, especially when more than 6 players are talking at once and the screen
scrolls quickly. When there is a particular goal or topic to the conversation, the spontaneous
mix of conversational threads (which is particularly conducive to textual play) can become
even more confusing. The conversants on TinyHall probably made use of the recognized real-
life conventions of hand-raising and moderation because they were attempting to discuss
one topic and wanted to maintain focus.

In sum, face-to-face models of turn-taking don't apply well to MUD conversation. In
MUD conversation, speakers make use of various textual devices for repairs and establishing
context and addressee, in the absence of visual or auditory channels. Conversation may be
multi-threaded and many types of floor structure are possible.

5.4 What are Back Channels?

In this section I discuss the type and frequency of back channels in the text-based MUD
discourse. Gumperz (1982) reports that back channels represent "one common way in which
conversational cooperation is communicated and monitored," and may include nods or other
body movements, or interjections like "ok," "aha," "right." Study of email messages (Black
et al. 1983) has found back channels to be significantly missing; study of ISDN half-duplex
video connection interactions has similarly found far fewer back channels than found in
face-to-face or full duplex video (O’Conaill et al. 1993). I maintain that back channels
do occur in MUD discourse (as Type 2 emotes) and are indeed important for determining
the attention state of an interlocutor, as well as establishing whether speaker intentions
have been understood. In a text-based medium where no physical cues are available, and
interlocutors may be called away from their terminal at any moment, these are particularly
important; back channel emotes and "says" play a large role in establishing achievement
of mutual understanding and facilitating a sense of co-presence (cf. McCarthy et al. 1993,
where co-presence was hard to achieve).

Schegloff (1982) reviews the literature on back channels briefly (p. 77):

The most common term now in use for such items, "back channel communica-
tion," was introduced by Yngve (1970), and includes a much broader range of
utterance types, including much longer stretches of talk. The term “back channel” has been adopted by Duncan and his associates (for example, Duncan and Fiske 1977), together with the broadened definition of the class. Duncan and Fiske (201-202) include not only expressions such as “uh huh,” “yeah,” and the like, but also completions by a recipient of sentences begun by another, requests for clarification, “brief restatements” of something just said by another, and “head nods and shakes.”

In this literature, they have been variously claimed to be signals of ongoing attention (Fries 1952, p. 49), to signal attention and understanding of what is being said (George don 1967, p. 44), and “to provide the auditor with a means for participating actively in the conversation, thus facilitating the general coordination of action by both participants” (Duncan & Fiske 1977, pp. 202-203).

In face-to-face conversation, Schegloff points out that back channels occur at possible turn-exchange points, until the speaker is obviously done and needing some other response. He considers them “continuers,” since they are abdicators of the turn exchange that otherwise might occur; they signal from a listener that a speaker may continue with an extended discourse structure, which presumably a listener must recognize is in progress. (Example 26 is from Schegloff 1982, p. 82.)

Example (26):

1 B: I’ve listen to all the things that chu’ve said, an’ I agree with you so much.
2 B: Now,
3 B: I wanna ask you something,
4 B: I wrote a letter.
   (pause)
5 A: Mh hm,
6 B: T’the governor.
7 A: Mh hm::,
8 B: -telling ’im what I thought about i(hh)m!
9 A: (Sh:::!) 
10 B: Will I get an answer d’you think,
11 A: Yes,

Note that after line 4, Speaker B paused, and a back channel was produced, enabling her to continue. At line 10, Speaker A is finally ratified as the next speaker.

The “confirmation feedback” discussed in Oviatt & Cohen (1988), important in task-oriented phone conversations to communicate understanding of goals, would also be considered back channels under most definitions. Grosz & Sidner (1986) discuss them within the larger category of “cue words” showing discourse structure understanding. Grosz & Sidner also looked at task-oriented dialogues; example 27 is an example from an expert-apprentice
dialogue (Grosz & Sidner 1986). The “ok” in line 5 below is considered a cue word, as is the “now” in line 6, which would not be considered a back channel by any definition.

Example (27):

1 A: One bolt is stuck. I’m trying to use both the pliers and the wrench to get it unstuck, but I haven’t had much luck.
2 E: Don’t use pliers. Show me what you are doing.
3 A: I’m pointing at the bolts.
4 E: Show me the 1/2" combination wrench, please.
5 A: OK.
6 E: Good, now show me the 1/2" box wrench.
7 A: I already got it loosened.

Schegloff (1982) focuses on nonspeech sounds like “mm hmm,” which he claims are used instead of longer forms like “I am listening” which would be more turn-like. However, they are conventionalized sounds that are well-recognized by a community. The frequency and type of back channels across cultures and across speech events within a culture may differ, often dramatically. White (1989) discusses the greater frequency of back channels in Japanese speech, for example. Debates, class lectures, and ceremonies differ in their turn-taking structures and therefore in their back channel type and frequency from those in dyadic conversations, which I will focus on now.

Duncan (1973) uses the term “back channel” to include non-turn material which shows the speaker how his turn is progressing for his listener, but he also includes “sentence completions,” “requests for clarification,” and “brief restatements,” which for others would be full turns (e.g., Duncan & Niederehe 1974). In Yngve’s (1970) original discussion of back channels, he makes clear that there are different types of back channels, along a continuum, from nonspoken attentive gaze, to nods, to murmurs, to short questions, to sections of talk as long as “filling in needed personal background so that the person having the floor could continue” (p. 574). He thus distinguishes between turns and “having the floor,” and considers much of back channel activity to consist of turn taking.

As pointed out in Coultlard (1977) and Goodwin (1981), there are differences of opinion over what constitutes a turn. Nods and other gestures are not usually considered turns in their own right, but murmurs of assent, or noises of the sorts Schegloff (1982) discusses are more questionable. As noted, for Schegloff (1982) it is important that the set of utterances examined are not full turns. Longer utterances (“I understand,” “I am listening”) could perform some of the same functions, but he maintains that the short forms are used because they are not full turns. The nonlexical status of the sounds he examines also prevents them from being heard as turns, despite their well-recognized conventional forms.
5.4.1 Function of Back Channels in the MUD

In this section I examine the types of back channels and their functions in the MUD register for ElseMOO.

In a MUD, back channels are full turns, since they are utterances like any other, and there is no “background” effect. As said earlier, Yngve’s (1970) and Edelsky’s (1993) distinction between “turn-taking” and “having the floor” is the correct one for the MUD discourse: a speaker may “have the floor” in some topic-controlling fashion, and choose utterance and turn size according to other factors, like how crowded the room is and how much interleaved talk there is. A speaker might similarly choose to provide feedback to another speaker, choosing not to change the topic.

In the discourse of the MOO, the routine utterances that I will call back channels are slightly different from those found in face-to-face conversation: they include some nonlexical imitations of speech sounds or laughter (“Tom hehs”), and some lexical descriptions of behaviors that are back channels “in real life,” like “lynn nods.” They also include conventional misspellings of other conventional forms (the back channel “hsm” came from a typo of the back channel “hms”) and other conventional typing shortcuts (e.g., “oic” comes from “oh, I see”).

Examples of back channel use in a MUD conversation occur in lines 3, 9, and 13 below in example 28.

Example (28):

1 Tom says, "only in look_self"
2 Karen says, "cool"
3 Karen nods.
4 Karen says, "oh"
5 Karen says, "there was another reason"
6 Karen says, "j----"j desc
7 Karen says, "huh"
8 Karen says, "now i can’t"

In this example, Karen is trying to describe how she wants some text to be laid out: lines 10-12 are her attempt to graphically represent the fields she wants, which consist of a name, a line of hyphens, and a description underneath it. (In line 11, her name was appended by the MOO server to the line with hyphens to indicate that she was the author of that “utterance.”)
Conventional expressions of puzzlement, as in line 9, are examples of "other-initiated repair" as discussed in Schegloff (1982). Since Schegloff suggests they occur at the same points in conversational interaction that back channel utterances might occur, I class them together with the other utterances I consider here (although Schegloff's analysis is quite different ultimately, because he is not interested in generalizing a single category). Longer, less conventionalized examples of repair initiation are not included, however (contra Duncan 1973 and Duncan & Fiske 1977; see section 5.3.2), because of difficulties with coding and searching for them automatically in the data. Affect signals or assessments like laughter, or the textual expression of it, are also considered to be in the supercategory of back channels for purposes of this chapter, for the same reason that short other-initiated repairs are. Additionally, I believe that affect responses are functioning very similarly in MUD discourse to the other types of back channel responses: they show attention and comprehension by registering appropriate emotional responses.5

Explicit back channels are used with great frequency among the ElseMOO population. Some of the more frequently used utterances have been encoded in easy-to-type commands that both document and encourage their common use. In particular, the Carpal Tunnel Syndrome feature command set (so-named because of several players' repetitive stress injuries, which inspired this collection of shorthand responses) is regularly used on EM; it makes available several of the regularly used back channels. For instance, if I type the "nd" command from the CTS command set, everyone in the room with me sees the output "lynn nods," which often indicates that I am attending or understanding what has been said to me. Other utterances in common use on the CTS command set include: "gg" which outputs "lynn giggles" (if typed by me), "h" for "lynn hehs" (simulation of laughter), "gr" for "lynn grins," "sm" for "lynn smiles," "/" for "lynn says," "?" (which approximates a questioning look, according to users' interpretations). The CTS commands set was originally created by two users about a year ago, but is now used by approximately one third of the regular population; the remainder use emotes for these routine utterances, e.g., by typing "::nods" at the prompt.

For this section, the back channels I considered in the MOO discourse are in the table below; all are conventional forms used by the community.

Example (29):
Back channels: nods, hsm, hms, hmm, hrn, oh, oic, ok, ah, yeah, yes, ?, giggles, laughs, grins, smiles, hehs

(In a later section, I consider some particular subcategories of back channels.)

Some of these occur as either emotes or "says," e.g., "Tom hmm" or "Tom says, "hmm," an alternation may be due to their non-lexicality. ("Tom says, "nod" would normally

5Note that line 5's "oh" is not a response to a remark from Tom, but is volunteered independently, indicating Karen has thought of something that she wants to discuss. This would not be considered a back channel, therefore (see Heritage 1984). My automated statistic generation process is not sensitive to such differences in use, unfortunately, and would classify it as a back channel.
be inappropriate, however, and gets classified as a “mav” by the community, so-called for a character on another MUD who regularly made modality errors in conversation; Smith 1995. However, I have seen it done consistently by players from Seashore MUSH.) Graphs of the number of emoted versus “said” back channels in four conversations of different lengths between two people each are shown in Figure 5.4. Note that in some cases the number of emoted back channels is almost equal to the number of “said” back channels. This result differs from the finding in Black et al. (1983) for electronic mail, where there were few back channels, due to lack of real-time responses. Black et al. (1983) claim “the conversational ‘back channeling’ associated with verbal discourse is minimal in discussions which use computer terminals, creating ‘dangling conversations’” (p. 62). MUD conversations do not “dangle” the way email or computer messaging (Murray 1989) conversations do, and clearly differentiating between synchronous, small-grained communication by terminal versus asynchronous larger-grained communication is necessary.

Despite problems with turn-taking models as a predictor of their placement in MUD discourse, there are clearly some regularities in back channel use in MUDs. In this rather focused discourse in example 30, notice that Karen and Tom produce them at the same points: in lines 4-5, after a point has been made, and then again at lines 14-15, after an expansion on another point. I produce one at line 11, agreeing with or acknowledging Tom’s comment in line 9.

Example (30):

1. lynn says, "cuts are followable in films, but I started wondering about them in the context of muds and teleporting"
2. lynn says, "without visual links, you lose relationship between spaces"
4. Tom says, "hmmmm"
5. lynn says, "and it could be done badly in a cd rom game too, no doubt"
6. Karen hears the mop, is so happy
7. Tom found some of the cuts in myst confusing.
8. lynn nods.
9. "but it’s still so slooooooowwww even with them"
10. Tom nods.

Schegloff (1982) points out that back channels are normally taken to indicate agreement, among other things. He suggests that they imply agreement because they occur in the same places as other-initiated repairs, yet they aren’t repair attempts. Usage as continuers in an ongoing discourse, passing up the opportunity for repair, leads to an implication of listener agreement. Clark and Schaefer (1989) similarly propose that they behave as acknowledgments or agreement markers.
Figure 5.4: Ratios of Emotes, Says, Emoted Back Channels, and Said Back Channels.
CHAPTER 5. MEDIUM EFFECTS: TURN-TAKING AND BACK CHANNELS 236

Back channels have been argued to show attention as well (e.g., Yngve 1970). Schegloff points out that back channels are at best “claims” of attention, which is different from “evidencing” it. This distinction may be important in a MUD, particularly where people are working while talking, and cannot always give their full attention to the text window. The persistent nature of the text, as well as the non-intrusiveness of it, allow for more casual involvement with it. Users can delay their experience of the conversation if they need to read it later in their buffer, something not possible with real-time audio or video, where a speaker expects more or less instant feedback. Full attention of the sort required in most face-to-face conversations is therefore not required in the MUD.

In Figure 5.5, I plotted the course of two sides of one long conversation on June 20, 1994. The left axis measures the number of utterances in each time period (since the last period) as indicated with the solid line, and the rate of speech, as calculated by words/utterances in the period, indicated by the dotted line. When a speaker is relatively inactive in conversation, there is a low utterance count (see time 43 minutes). A higher words/UTS rate indicates longer utterances are being produced (see time 10 minutes on graph lynn talking to Tom). Time increments are about 11 minutes. The star symbols show number of back channels received from the interlocutor during each time period, measured along the right axis. The graph of lynn talking to Tom has Tom’s back channels plotted over it, in other words.

In their analysis of task-oriented conversation, Gross and Sidner (1986) found that cue words like “ok” and “yeah” showed comprehension of a speaker’s goals had been achieved; Schegloff (1982) suggests that back channels are often “continuers,” abdications of a full turn from a hearer, which essentially give permission to a speaker to continue developing a complex discourse structure like a narrative. If back channels were only functioning as an indicator of comprehension of a speaker’s plans or response to the development of an extended discourse structure, there should be fewer back channels during periods with little interaction. However, periods with low utterance counts for both conversants (43 minutes, 78 minutes) nevertheless show the presence of several back channels from the interlocutors. (Even the inclusion of some repair and affect responses in my counts ought not to prejudice the results; especially if both those categories of response are similarly associated with complex discourse structures; see below.) The appearance of these signals in such periods probably indicates that a potential interlocutor is attending and may be available for more extended conversation.

It is difficult if not impossible, however, to totally distinguish between “acknowledgment” functions (Clark and Schaefer 1989) which indicate understanding, and “attention” functions alone, since the attention markers are always occurring in response to utterances and may perform both functions. In the next section I break down back channels by type more thoroughly, but each may still perform multiple or ambiguous functions. The lack of correlation between back channels and periods of increased utterance rate, however, suggests that the conclusion that they function as “continuers” when a speaker is producing an extended turn at talk is not the whole story, at any rate. The conversants in this phatic, largely non-task-oriented conversation are in a “continual state of incipient conversation” (Schegloff & Sacks 1974), analogous to that achieved by the two-party situations Goffman
Figure 5.5: Graph of Rates of Utterance Overlaid with Back Channels
(1963) describes: “communication arrangements that seem to lie halfway between mere co-presence and full-scale co-participation.” I provide several examples of back channels during these periods of low participation below.

In the following excerpt from the period around the 43 minute mark, the conversation has moved off the previous topic which was a somewhat tense one, and the interlocutors are registering their continued alertness, even while they document their actions in real life. In lines 1-2 of example 31, I illustrate that I am reading email in another window, and paste a section from one message. After a desultory exchange on that topic, including a back channel at line 6 indicating comprehension, Tom begins playing with names in thought bubbles (lines 7-9) and then reports singing, a common practice while listening to music and MUDding at the same time. I respond with back channel responses in lines 12 and 14, initially indicating I am still alert, and then acknowledging receipt of information, before making another desultory conversational offer in line 15.

Example (31):

1 lynn sees OJ all over the popcult list, of course.
2----------------------------------lynn------
3 In any case, thanks to OJ, Al, and the LA chopper teams and reporters for providing all of us cult stud folx with yet another a perfect Baudrillardian moment...
4---------------------------------- lynn stops pasting--
5 Done @pasting.
6 Tom says, "al?"
7 lynn dummo.
8 lynn says, "media coverage somehow."
9 Tom says, "ah"
10 Tom . o 0 ( oj et al )
11 Tom . o 0 ( woj simpson )
12 Tom . o 0 ( homer j simpson )
13 Tom [sings]: who threwed lye on my dog?
14 Tom wonders if he could fall asleep.
15 lynn says, "hmm. go home and sleep."
16 Tom is home.
17 lynn says, "oh."
18 lynn says, "I wonder why I keep dreaming about food."

Notice line 12’s “hmm” is part of a larger turn. MUD back channels do not have to be separate turns, and as we saw, turn status is hard to define in a MUD.

Narrative emotes of what is happening “in real life” or elsewhere are common during
semi-idle periods (Type 4, from chapter 4). Narrative emotes tell the listener that attention may be split and how. (See Figure 4.1; narrative emotes are slightly more common in group conversation than in dyadic conversation on the whole.) Here is another example from around the 43 minute time: note that lines 5-6 are a split turn.

Example (32):

1 Tom [sings]: i live
2 Tom [sings]: where it's
3 Tom [sings]: graaaaayye
4 lynn gets another random guest 'hi' on lm.
5 Tom says, "hm"
6 Tom says, "do you answer these?"
7 lynn says, "the other day I did, but I am not today."

The distribution of back channels during periods of mutual rapid, dense conversation suggests that some of their functions in periods with less conversation (marking attention, showing understanding or confusion, providing assessment like laughter) are being taken over by other types of utterances (or by longer phrases that I am not counting in my scripts as back channels). The number of back channels given by speaker A to an interlocutor B generally increases when B's number of utterances increases, but often stays low if instead A's utterance rate increases in parallel (see time 21 minutes and 63 minutes on Figure 5.5). If A's rate increases too, A is not being a passive listener, but is involved actively in the conversation as well, which may result in fewer back channels. In example (33), from around the 60 minutes mark in Figure 5.5, the confusions are cleared up with explicit questions, rather than conventionalized shorthand expressions like the single question mark or "huh." (The "ow ow ow" utterance in line 5 means that it's harder to type comfortably at home.)

Example (33):

1 Tom thinks about when he'd wake up if he went to sleep now, and whether it'd be safe to bike.
2 lynn says, "bike where?"
3 Tom says, "work"
4 lynn says, "how hard is it to work from home?"
5 Tom says, "ow ow ow"
6 Tom says, "plus, i can't do any interesting jupiter stuff"

In sum, there are frequent back channels in periods of only desultory interaction, and the back channels appear to function generally as evidence of continued attention, necessary in a text-based medium where an interlocutor may be invisibly called away or not attending.
They may decrease when more substantive means of registering attention and understanding are used, in periods of more active participation.

5.4.2 Situated Back Channel Functions

As noted in the conversation analysis literature, different back channel utterances may perform slightly different functions, and back channels may behave differently in different contexts. Jefferson (1981) suggests that "mm hmm" is a passive recipient token, while "yeah" implies its producer may soon take the floor. The "hmm" utterance or emote (and variants "hsm," "hms," and "hm") in the MOO is not equivalent to the "mm hmm" utterance; it probably functions similar to "hmm" in some "real life" conversation, which represents a sign of thought or discomfort with an interlocutor's previous statements, carrying a suggestion that an explanation of the cause of discomfort or other further comment will be forthcoming. In Examples 31 and 32, "hmm" was accompanied by further comment. In Example 30, while Tom did not respond further immediately after his "hmm," he was the next speaker on the topic. In a non-response situation, "hmm" (or this variant, the tensed "hms") functions as a sign of internal disquiet:

Example (34):

Honda hms, that URL doesn't seem to work...

If there is no accompanier to "hmm," a question may be forthcoming from an interlocutor. In this example, I said "hmm" about the guest's wandering behavior. Tom didn't make the connection and asked about it. (Of course, the specific function of the back channel as representing disquiet is not established by this example; it may simply be that its nature as a back channel requires that there be an item to be responded to, which was not clear in this non-conversational example.)

Example (35):

[at 4:00 P.M.]
A guest scrambles up from the ditch at the east side of the road.
The guest descends into the sewer.
A guest emerges from the drain.
The guest scrambles down into the ditch to the east.
[at 4:01 P.M.]
lynn hmm
Tom says, "?"
lynn says, "just that guest being so crazed"
Tom nods.

The "nod" may either function passively as a token of understanding or agreement, with no further comment expected; or it may perform as a required turn, indicating "yes,"
in which case it is not precisely a back channel (although my automated search code will classify it as one). The cases in Example (36) contrast with the uses of nods in Example 30, where they were not necessary responses, but signalled continued attention and/or agreement. (Note that in “real life” conversation, a nod without a vocal acknowledgment in this situation would be rude at best).

Example (36):

Ray says, "incidentally, ny works multicast
          on linux"
Ray tested last night
Honda [to Ray]: That’s video stuff?
Ray nods
Honda [to Ray]: Jubilee Frederick’s package?
Ray nods
Honda [to Ray]: Cool ...

These functions of nod responses as full-turns (here, second pair parts in adjacency pairs) highlight their underspecified, interactional meaning. In periods of little conversational action, they are indicators of attention, with some suggestion of agreement; in answer to questions, the agreement meaning becomes prominent and they can perform a turn. This is consistent with Clark and Schaefer (1989), where back channel utterances function minimally as indicators of attention, and may also be acknowledgments (which are continuers, in Schegloff 1982’s sense), or even more significant contributions. Gricean cooperative principles help make the interpretation of the nod into an appropriate response to a question.

In the table below, I outline an expanded and further broken down categorization of utterances that function as back channels, affect, or repair initiators. I believe it is difficult if not impossible to separate affect out from the back channel function in this medium, since an appropriate emotional response to a turn (e.g., a laugh) indicates both attention and understanding just as well as a nod does. Furthermore, the distinction Goodwin (1985) makes between continuers and assessments—that the former occur between turns, and the latter overlap turns—is not available in the MUD medium, where overlap cannot occur. However, I have here labelled them as different types. “Status” indicators are basically the explicit acknowledgment utterances like “yeah” and other-initiated repair utterances, which may be emotet or “said.”

Example (37):
<table>
<thead>
<tr>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Affect| lol gol grin smile laugh heh  
|       | giggle hee frown sigh eyes |
| Nod   | nod                       |
| Status| hsm hmm hrn ? ok yeah  
|       | yes uh um err hms         |

Figure 5.6 illustrates breakdowns of the number of each type in 20 minute conversations for dyads and groups, two conversations of each type. It appears predictably that affect responses increase in large group conversation, while the number of status utterances decreases. This is probably a function of the party atmosphere, in which it is harder to have a focused conversation and account for misunderstandings in a group.

A representative slice of party conversation with affect back channels is below. Note how many parallel conversations are occurring, too. (In line 24, “gol” stands for “giggles out loud.”)

Example (38):

1. Tom is simply ignoring George now that he’s still held at arm’s length.
2. Freddie finally got his pepsi
3. Freddie says, "well slice actually"
4. Karen [to Shelley]: yay
5. Karen . o 0 ( orange sherbert )
6. Freddie [to Shelley]: but you haven’t read them yet
7. Shelley [to Freddie]: then how do i know they’re good?
8. George whuggles Tom.
9. Freddie [to Shelley]: how do you know they’re good?
10. lynn says, "awww."
11. Freddie nopds
12. Tom whuggles George.
13. Tom grins.
14. Bonnie blinks
15. Freddie noocoods
16. Shelley [to Freddie]: cuz i read em, doof
17. Shelley giggles.
18. Bonnie hehs, gets nominated for ARB
21. Tom HASSLED GEORGE INTO WHUGGLY SUBMISShm, that sounds a little weird
22. Freddie [to Bonnie]: did you really?
23 Karen eyes Tom warily.
24 George grog
25 Bonnie nods to Freddie.
26 Bonnie says, "bizarre"
27 Karen: o o (too weird)
28 George mmm, ARB
29 Freddie says, "wow... everybody is"
30 Tom: You have new mail (172) from ARB Petition Core Object
          (#5250).
31 Bonnie nods.
32 Karen [to Freddie]: what, weird?

Although not directly comparable, O’Conaill et al. (1993) found an average of 7 back
channels in 20 minutes of ISDN video, as compared to 30.5 in LIVENET video and 60.8 in
face-to-face conversation. Classification and identification of back channels in those media
are a very different prospect from the textual interactions my Figure 5.6 represents, however.

In summary, different back channels may perform different functions, highly contextually
determined. In the MUD medium where no overlap is possible, differences between
continuers and assessments may not be possible. Affect signals function as back channels;
in large party situations, there are more affect utterances than in dyadic conversations and
there are fewer "status" markers of (mis)understanding.

5.5 Conclusion

Consideration of turn-taking and back channel usage are important in analysis of the register
features of the MUD discourse, since they are critical aspects of interaction in the medium
as a whole.

In this chapter, I discussed how face-to-face models of turn-taking do not map well
onto MUD turn-taking, because of the lack of simultaneity in the medium. Overlap and
hence interruption do not occur in the way that they do in face-to-face conversation. Since
speakers do not need to negotiate for the turn in MUD discourse, but rather for possession
of the floor in a more topic-focused manner, theories of floor are more fruitfully considered.
The notion of the collaborative floor is particularly useful, since multiple topics are often
handled in parallel in the discourse of the MUD. Furthermore, I showed that turn-size is
generally shorter than in face-to-face conditions, contrary to predictions in the literature.
This finding supports the Rational Actor Hypothesis, that users adapt to the medium and
learn to produce smaller sized turns in order to maintain interactivity and a sense of co-
presence.

Finally, I suggested that the general distribution of back channels in several MUD con-
versations shows that they are used for maintaining a sense of co-presence and awareness
in a conversation, not just for signaling comprehension, assessment, or recognition that a
Figure 5.6: Types of Back Channels in Dyadic and Group Conversation
complex discourse structure is under construction. Although each back channel response may perform slightly different functions, as a class they share some characteristics: they are conventional, often non-lexical, responses used by the community in similar places in the conversational interaction. Some of them describe non-linguistic actions like nods or laughter. MUD conversation does not map well onto turn-taking models of face-to-face conversation, since turn boundaries are problematic to define, so MUD back channel placement is not easily described in terms of transition relevance points. However, their overall distribution suggests that they increase when an interlocutor’s rate of speech increases, unless both speakers become very active in parallel. Otherwise, their occurrence in periods with relatively low utterance rates suggests they help maintain a sense of continued conversational context and co-presence even when focused, topic-driven conversation is lacking.