Speaking Through Text: The influence of real-time text on discourse and usability in IM

Jacob Solomon, Mark Newman, Stephanie Teasley
School of Information, University of Michigan
105 South State St.
Ann Arbor, MI 48109-1285
{jacobbs, mwnewman, steasley}@umich.edu

ABSTRACT
Real-time, character-by-character transmission of messages in synchronous forms of text-based communications has seen a recent resurgence in CMC. We evaluated the impact of real-time text display on the usability of an instant messaging (IM) client. Participants were randomly assigned to dyads to participate in two discussion tasks using IM with both real-time text and enhanced message-by-message display (i.e., line-by-line display with additional cues to show when the remote party is typing). We found that real-time text helped users better coordinate turns and lead to less self-editing of messages, but had no overall influence on users' typing ability and provided minimal support for collaborative completion of sentences. Users who typed less or had less experience with IM tended to prefer real-time text. These findings have significance for several forms text-based CMC, including IM, chat, text telephony, and collaborative document editing.

Categories and Subject Descriptors
H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces – *computer supported cooperative work, synchronous interaction.* H.4.3 [Information Interfaces and Presentation]: Communications Applications, *computer conferencing, teleconferencing, and videoconferencing.*

General Terms
Human Factors.

Keywords
Real-Time Text, Instant Messaging, Turn-Taking

1. INTRODUCTION
Instant messaging (IM) and chat applications have varied in their approach to displaying short, text-based messages to interlocutors. Internet-relay chat (IRC) and other early protocols required users to compose a message privately, transmitting only the completed message by using the "Return" key. Other interfaces, such as Unix Talk and ICQ, remove the private composition space by transmitting each character of a message in real-time as it is typed. This approach is frequently known as real-time text.

Previous research comparing the two styles of text display ([5], [6], [7], [10], [18]) revealed advantages to each approach. Eventually, industry largely abandoned real-time text display and adopted an augmented form of the message-by-message approach in which ambient information about keyboard activity is displayed (i.e., *User is typing…*). This form of display, which we will call "enhanced message-by-message display," has been included in most IM clients for the past decade; however, there is a lack of research comparing it directly to real-time text.

New applications have emerged for which real-time text has apparent value. Accessibility groups ([11], [12]) have advocated the implementation of real-time text for text telephones and IM clients, with the intent of making a more conversational style of discourse accessible through mediated technologies. Tools such as Google Wave[1] and Etherpad[2] have also implemented real-time text in communication and collaborative document editing. In many of these applications, real-time text has been implemented due to a desire to better simulate spoken conversation through text. With the growing popularity of new applications of real-time text, it is important to re-evaluate its relative advantages and disadvantages compared to enhanced message-by-message text display. The design of tools for text-based CMC can be better informed by an understanding of how real-time text may influence user behavior and the user experience.

Voida, Newstetter, and Mynatt [17] point out that "it is often ambiguous whether use of instant messaging aligns with conventions of written communication, verbal communication, or exists somewhere between the two." Real-time text may improve IM by clarifying this ambiguity and facilitating a speech-like style of communication. In this project we investigate the extent to which real-time text simulates spoken dialogue, and look at the effect this has on the usability of a CMC tool.

In this paper, we report the results of an experiment in which we compared real-time text to enhanced message-by-message display in IM in terms of the impact on communication and overall

GROUP'10, November 7–10, 2010, Sanibel Island, Florida, USA.

[1] http://wave.google.com/ (retrieved May 19, 2010)

[2] http://etherpad.org/ (retrieved May 19, 2010)

usability. Specifically, we looked at the coordination of turn-taking, self-editing, typing performance, collaborative completion, and subjective preference, as well as overall affect for each style of text display. We found that real-time text display facilitated communication by helping users coordinate turn-taking. One reason for this improvement in turn-taking is that real time text lead to less self-editing, perhaps allowing users to remain more engaged in the conversation. In contrast to our expectations, we also found that real-time text in an IM client offered only minimal support for collaborative completion of a sentence across turns. Contrary to earlier findings by Vronay, Smith, and Drucker [18] real-time text did not appear to influence users' typing ability. Across users overall, it was neither more or less preferable than message-by-message display, although we found specific factors which influenced preference in either direction, such as typical IM usage by participants and the ratio of "spoken" to "listened" turns in the IM conversation

2. RELATED WORK

There has been no work comparing real-time text to enhanced message-by-message display, even though the former style is now nearly ubiquitous in IM clients. However, there is prior work comparing real-time text to message-by-message text-only display (without keyboard status). Turn-taking has been the most widely investigated factor in the research on real-time text. Several studies ([5], [7], [10], [18]) have looked at turn-taking in chat or IM and compared real-time text to a message-by-message style of display and demonstrated that real-time text serves to help users take turns. In these studies, the framework introduced by Sacks, Schegloff, and Jefferson [14] was largely used for defining turn-taking as adjacency pairing, in which a speaker's turn appears adjacent to the most relevant turn from another speaker (i.e. a question is followed by an answer to that question, a statement is followed by a response, or a greeting is followed by another greeting).

Campbell [1] took a different approach to turn-taking and looked at the level of simultaneous typing in modern IM clients with an enhanced message-by-message style. He found that IM users type simultaneously about 17% of the time. This can be compared to the 5% rate observed in face-to-face conversation [8] and the 8% rate observed in telephone conversation [15]

Vronay et al. [18] also found that real-time text led to improved typing ability, although it also led to user anxiety over displaying their own typing ability. This is consistent with what Voida et al. [17] describe as an "articulateness tension," in which users try hard to present themselves as good typists and use proper grammar. This tension has been observed in other research on IM user behavior ([2], [9])

3. STUDY GOALS

In our study, we conducted an experiment to understand whether real-time text offers a better alternative than enhanced message-by-message text display. We investigated the style of discourse in the two display conditions using analyses from research on spoken discourse. As turn-taking contributes to grounding within a conversation [3], we evaluated turn-taking by looking at adjacency pairing and simultaneous typing. In addition, we looked for evidence of collaborative completion (i.e. finishing each others' utterances) in the real-time text display condition, as this is a characteristic of spoken conversation and has been shown to positively benefit CMC [13]. We asked whether real-time text would support collaborative completion by allowing users to finish incoming utterances, by either anticipating incoming content within a partner's turn messages or by completing the partner's sentence in their next turn.

We also examined whether real-time text display would influence the overall usability of IM. Specifically, we were interested to see if real-time text would increase the "articulateness tension" described by Voida et al. [17]. To do so, we tried to replicate Vronay et al.'s [18] finding that real-time text influences typing ability. In addition, we looked at how real-time text would influence self-editing of messages. Finally, we examined how users feel about real-time text display, and investigated factors that influence preference for or against real-time text.

4. METHODOLOGY

4.1 Participants

Twenty-four native English speakers with experience using IM (14 male and 10 female) were recruited from the campus of our university through flyers advertising a study on instant messaging.

4.2 Procedure

All participants were randomly assigned to dyads and participated in two different discussion tasks using both forms of text-display. They were randomized across their first study condition (real-time text or message-by-message) and task order was balanced by condition. For one task we asked participants to discuss movies and supplies they would bring to a weekend cabin retreat. For the other task, we asked participants to determine the three most critical transportation issues on our university's campus and develop a plan for investigating them. Fifteen minutes was allotted for each discussion. AIM 7.0 was used, which allowed real-time text to be turned on and off without any other interface changes. Participants were not told which other participant in the session was their partner.

Between tasks, participants were given the IBM Computer System Usability Questionnaire (CSUQ) and asked about their typical IM usage on a five-point scale ranging from "Infrequently" to "5+ hours a week." Participants later answered a free-response post-test questionnaire about their preference of text display. This post-test questionnaire asked participants to describe instances of collaborative completion and any resulting problems caused by incorrectly completing incoming messages.

4.3 Text Analysis

Chat logs, screen recordings, and keystroke logs were used to collect data from the first ten minutes of each session (the period in which all conversations were active). Each conversation was coded and broken into adjacency pairs. Clark and Schaeffer's [4] description of adjacency pair types was used to code the conversations. A turn (text delimited by the "Return" key) was

U01(5:38:23 PM):	do you like action movies?	pair 1 part 1
U02(5:38:24 PM):	so sleepy hollow and amityville horor	pair 2 part 1
U02 (5:38:26 PM):	yeah	pair 1 part 2

Figure 1. A disrupted adjacency pair.

coded as a first pair part or a second pair part, and pairs that belong together were identified. When the two parts of an adjacency pair were separated by a part of a different adjacency pair, it was coded as a disrupted pair. Figure 1 demonstrates an example of a disrupted adjacency pair. The disrupting part could come from either speaker. A disrupted pair could result from a speaker asking a question, then asking another question before their partner submitted a response to the first question. It could also occur if participants merely tried to start new adjacency pairs nearly simultaneously. It could not result from an "insertion sequence" [8] such as following a question with a clarifying question.

Self-editing was measured by the number of sequences of one or more consecutive "Backspace" keystrokes.

Stopping points for each turn were recorded from the keystroke logs. Adjustments were made to "false start" turns where both partners began typing simultaneously but one quickly conceded the turn space by deleting their entry. The adjusted stopping points were used to measure the amount of simultaneous typing (measured in seconds) and typing ability. Typing speed was measured as keystrokes per second (KSPS) and accuracy measured as keystrokes per character (KSPC), as suggested by Soukoreff and MacKenzie [16].

The number of turns, words, and characters was recorded for each dyad and for each participant.

Paired t-tests compared the means of these quantitative measures between conditions. A Mann-Whitney U-test compared the CSUQ Likert scale responses. The post-test questionnaires and CSUQ free response sections were analyzed qualitatively. Logistic regression was used to find factors which contributed to preference for one style of text display over the other.

5. HYPOTHESES

The study investigated six hypotheses.

H1: Real-time text will lead to fewer disrupted adjacency pairs than enhanced message-by-message IM.

H2: Real-time text will lead to less simultaneous typing than enhanced message-by-message IM.

H3: Users will engage in collaborative completion, as either anticipated (as reported in the post-test questionnaire) or visible in the message stream.

H4: Real-time text will lead to more self-editing as a result of users beginning to type responses based on an incorrect anticipated completion of an incoming message.

H5: Users will type faster using real-time text.

H6: Users will type more accurately using real-time text.

6. RESULTS

Table 1 summarizes the two-tailed paired t-tests from the experiment. H1 and H2 were both confirmed, that is, dyads using real-time text had fewer disrupted adjacency pairs, t (11)=-4.11, p<.01, and less simultaneous typing, t (9)=-2.83, p<.05. H4 was not confirmed as users made fewer self-edits when using real-time text, t (21)=2.32,p<.05. Note that the keystroke files for two observations were corrupted and could not be analyzed. H5 and H6 were also not confirmed as no differences in either typing speed or accuracy were recorded.

Table 1. Quantitative results

	Real-time text	**Message-by-message**	**Paired t-test**
% of disrupted adjacency pairs	29% (11.3%)	39% (9.8%)	t(11)=-4.11; p<.01
Number of deletion sequences	38.4 (13.9)	51.8 (22.83)	t(21)=2.32; p<.05
% of time spent typing simultaneously	10.8% (8.5%)	17% (8.6%)	t(9)=-2.83; p<.05
Keystrokes per second	2.884	2.267	Not significant
Keystrokes per character	1.363	1.359	Not significant
Preference	10	14	Not significant (binomial distribution)

The qualitative data from the post-test questionnaire suggest that H3 was only partially confirmed. Participants reported that they did try to anticipate how messages would end but they felt it was impolite to begin responding until the incoming message was complete. These results suggest that collaborative completion is only minimally supported by real-time text in an IM client.

The Mann-Whitney U-test found no differences between text display conditions on any CSUQ questions, and neither system was overwhelmingly preferred. Logistic regression did, however, find that participants' typical level of IM usage influenced preference, with heavy IM users more likely to prefer message-by-message display (β=1.2293, p<.05). More interestingly, when using enhanced message-by-message display, participants who did more typing preferred enhanced message-by-message display and those who did more reading preferred real-time text display (β=.02209, p<.05), indicating that one's preference for text display style may be influenced by one's role in the conversation.

In the free-response portion of the CSUQ, 5 of the 12 real-time text respondents listed real-time text as one of the "most negative features" of AIM. One user noted "On AIM, it is annoying to see what someone is typing before they're done typing. I would rather see complete ideas."

7. DISCUSSION

Our data suggest that real-time text display does facilitate a *closer* simulation of spoken discourse than enhanced message-by-message display. Specifically, in the real-time text display condition we found fewer disrupted adjacency pairs, less simultaneous typing, and less self-editing, in addition to some support for collaborative completion of sentences. Despite these data, it is clear that real-time text falls short of offering a simulation of spoken conversation. The turn-taking system in spoken conversation [14] consists of numerous strategies for managing adjacency pairs, making disruption an infrequent occurrence in spoken conversation (when typical politeness norms are enacted). Yet nearly 30% of adjacency pairs displayed in our IM-based conversations were disrupted when using real-time text. The 10.8% of simultaneous "talking" was comparable to 8% of simultaneous talk that has been measured in telephone conversation, but it still represents a relative loss in turn-taking coordination. And participants still took some liberty to self-edit their messages, a practice which is less frequent in spoken discourse. And although collaborative completion was supported in a limited way, there was little evidence in the text that users can actually engage in this behavior explicitly. At best, real-time text offers a slight improvement over enhanced message-by-message display to some of the traditional problems with simulating discourse through text, but does not solve them completely.

These findings suggest other important usability considerations for the design of IM tools, and of other tools for text-based CMC. The relative reluctance to self-edit messages suggests that real-time text may complicate the "articulateness tension" by making edits and revisions to a message more visible. These data also suggest that there is a tradeoff between support for turn-taking and support for freedom to self-edit that should be considered in the design of CMC tools. Real-time text appears to support message consumers at the expense of message producers, as evidenced by both the self-editing data and the data regarding preference for style of text display.

As systems continue to be designed for simulating spoken conversation through text, designers should consider real-time text a useful but insufficient solution. Designers should look to compliment real-time text with new innovations which better support collaborative completion. Systems such as collaborative text editors, which may not have set such a simulation as a goal of the system, should consider the other usability issues presented by real-time text in light of these data.

8. REFERENCES

[1] Campbell, J. D. (2004). Instant messages: a framework for reading between the lines. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (pp. 519-522). Chicago, Illinois, USA: ACM.

[2] Campbell, J. D. (2005). Does spelling matter in instant messaging?: answers from measuring error correction frequency. In *CHI '05 extended abstracts on Human factors in computing systems* (pp. 1244-1247). Portland, OR, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=1056808.1056887&coll=GUIDE&dl=GUIDE&CFID=75544600&CFTOKEN=55411312

[3] Clark, H., & Brennan, S. (1991). Grounding in communication. In *Perspectives on socially shared cognition* (pp. 127-149). American Psychological Association.

[4] Clark, H. H., & Schaefer, E. F. (1989). Contributing to Discourse. *Cognitive Science*, *13*(2), 259-294. doi:10.1207/s15516709cog1302_7

[5] Garcia, A. C., & Jacobs, J. B. (1999). The Eyes of the Beholder: Understanding the Turn-Taking System in Quasi-Synchronous Computer-Mediated Communication. *Research on Language & Social Interaction*, *32*(4), 337. doi:10.1207/S15327973rls3204_2

[6] Hancock, J. (1999). The importance of coordination devices in text-based, task-oriented conversations. In *CHI '99 extended abstracts on Human factors in computing systems* (pp. 310-311). Pittsburgh, Pennsylvania: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=632905

[7] Hancock, J. T., & Dunham, P. J. (2001). Language Use in Computer-Mediated Communication: The Role of Coordination Devices. *Discourse Processes*, *31*(1), 91. doi:10.1207/S15326950dp3101_4

[8] Levinson, S. C. (1983). *Pragmatics*. Cambridge textbooks in linguistics. Cambridge [Cambridgeshire]: Cambridge University Press.

[9] Nardi, B. A., Whittaker, S., & Bradner, E. (2000). Interaction and outeraction: instant messaging in action. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work* (pp. 79-88). Philadelphia, Pennsylvania, United States: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=358975&dl=GUIDE&coll=GUIDE&CFID=80866997&CFTOKEN=89541859

[10] Phillips, B. (2000). Should we take turns?: a test of CMC turn-taking formats. In *CHI '00 extended abstracts on Human factors in computing systems* (pp. 341-342). The Hague, The Netherlands: ACM.

[11] Real-Time Text Taskforce. (n.d.). . Retrieved January 4, 2010, from http://www.realtimetext.org/index.php

[12] Real-time text versus other text based communications. (n.d.). . Retrieved September 14, 2009, from http://www.ictrnid.org.uk/whyrtt.html

[13] Roschelle, J., & Teasley, S. (1992). *Journal of the Learning Sciences*, *2*, 235-276.

[14] Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696-735.

[15] Shriberg, E., Stolcke, A., & Baron, D. (2001). Observations on Overlap: Findings and Implications for Automatic Processing of Multi-Party Conversation. *EUROSPEECH-2001*, 1359-1362.

[16] Soukoreff, R. W., & MacKenzie, I. S. (2003). Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 113-120). Ft. Lauderdale, Florida, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=642611.642632&type=series

[17] Voida, A., Newstetter, W. C., & Mynatt, E. D. (2002). When conventions collide: the tensions of instant messaging attributed. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* (pp. 187-194). Minneapolis, Minnesota, USA: ACM.

[18] Vronay, D., Smith, M., & Drucker, S. (1999). Alternative interfaces for chat. In *Proceedings of the 12th annual ACM symposium on User interface software and technology* (pp. 19-26). Asheville, North Carolina, United States: ACM.