Linguistic and Critical Analysis of Computer-Mediated Communication: Some Ethical and Scholarly Considerations

SUSAN HERRING

Program in Linguistics
University of Texas at Arlington
Arlington, Texas, USA

This essay compares two proposals (Cavazos, 1994; King, this issue) relating to whether and how CMC researchers should cite electronic messages used as data. Although the proposals prescribe opposite solutions, both contain similar assumptions about the nature of CMC (e.g., that it is homogeneous; that members of a "virtual community" have shared agendas) and about the nature of research (e.g., that it is content focused; that it is ideally consensual; that it should not affect the researched in any way). These assumptions are argued to reflect discipline-specific biases that exclude other legitimate forms of CMC research. Two examples are discussed of research paradigms that are excluded by the guidelines: linguistic analysis in the positivist tradition, and critical analysis in the social realist tradition. The critical paradigm in particular raises a number of additional ethical considerations not addressed by the proposed guidelines. It is suggested that existing ethical guidelines within each discipline largely suffice to guide on-line research, with the addition of a CMC-specific recommendation clarifying the rights and obligations of researcher and researched in restricted-access as compared with open-access on-line groups.

Keywords computer-mediated communication (CMC), ethics, guidelines, linguistics, pseudonyms, scholarship, social criticism

The Problem

In the early years of CMC research—which is to say, until very recently—those of us researching CMC (computer-mediated communication) had no choice but to make up rules and procedures as we went along. Quite simply, the phenomena we were investigating hadn't been in existence long enough yet for others to have paved the way with methodologies, ethical guidelines, and the like. Being among the first brought with it a sense of discovery and exhilaration, but also uncertainty at times on how best to proceed. Thus it was with a vague sense of relief that I welcomed the first suggestions of how to cite

Received 15 December 1995; accepted 30 January 1996.

Susan Herring is an Associate Professor of Linguistics at the University of Texas at Arlington. Since presenting the first results of her research on gender and CMC in 1992, she has published 12 articles and has delivered numerous talks on the topic, including in Europe and Asia. She is the editor of an interdisciplinary volume on CMC, and guest editor of a special journal issue on Computer-Mediated Discourse Analysis, both forthcoming in 1996. Her work has been cited in *Newsweek*, the *Boston Globe*, *Der Spiegel*, the *Digital Media Newsletter*, and in numerous regional newspapers. Areas of her CMC research include politeness and netiquette, schematic organization of electronic messages, and gender and rhetorical style.

Address correspondence to Susan Herring, Program in Linguistics, Box 19559, University of Texas at Arlington, Arlington, TX 76019, USA.

The Information Society, 12:153–168, 1996

0197-2243/96 $12.00 + .00

e-mail messages in scholarly publications, how to cite electronic journal publications on one's curriculum vitae, and advice on other such practical matters.

Still, there was as yet no conventionally agreed upon advice regarding whether (or when) one should use participants' real names when citing electronic messages as data. This is a matter of special concern to me as a linguist, since my research focuses on the language used in electronic interactions and involves quoting portions of actual messages verbatim. Early on, I made the decision to use pseudonyms or avoid mention of names rather than reveal the actual identities of my data sources, in part because this is the usual practice in linguistic research, and in part because my research is critical of the language patterns used by some participants. My intention was to respect the privacy of individual participants, while preserving the academic freedom to criticize. But is this defensible on ethical grounds?

Now, at last, two sets of proposals related to this issue have been publicly advanced. Unfortunately, however, the two proposals aim to establish guidelines that are mutually contradictory in crucial respects. Further, neither appears to have envisioned the possibility of CMC research that is linguistic in focus or critical in nature.

Two Proposals

The first proposal comes from legal scholars, most notably Edward Cavazos in a recent book entitled *Cyberspace and the Law: Your Rights and Duties in the On-Line World.* It asserts, in essence, that all messages posted via computer networks are published works and hence protected by copyright law. Quoting a message or part of a message in another published work without giving full credit to the source (naming the message writer, the group it was posted to, the time and date, etc.) is a violation of copyright and legally actionable. According to this view, one should use only participants' real names, and indeed provide further identifying particulars, whenever an electronic message is quoted.

The second proposal is presented by Storm King in his article in this issue. It asserts that all messages posted to computer discussion groups are potentially private in terms of how they are intended and perceived by participants within the groups. In order to protect the "perceived privacy" of participants in electronic interactions, no potentially identifying characteristics of the data should be reproduced in scholarly work, including participants' names, the time or date of the message, or the real name of the group itself. According to this view, one should paraphrase rather than quote messages verbatim, or if messages are quoted, they should be carefully expunged of all group- and author-specific information.

The contradiction between these two views is obvious—one says to reveal one's sources, the other to hide them, regardless, in both cases, of the particular circumstances of the communication. On the one hand, such absolutist positions might hardly seem worth serious discussion. It is untenable to claim that all CMC is copyrightable (some is trivial; consider, for example, the one-word message "Hi" posted by a participant on a chat channel), just as it is untenable to maintain that all CMC is private (consider an advertisement broadcast simultaneously—"spammed"—to thousands of newsgroups on the Usenet). Each of these views appears to assume one particular type of CMC (for example, open debate of intellectual issues in the first case, and self-revelation of sensitive information in self-help groups in the second) and to generalize recommendations based on that type to all of cyberspace. However, cyberspace is a vast and varied domain, and rules that seek to generalize indiscriminately across all varieties of CMC do not "fit" the nature of the phenomenon.

Less obvious but equally misguided, each of the proposed guidelines assumes a particular model of scholarship and extends it to all CMC research. The idea that the source of all electronic messages should be credited assumes that (1) the messages are cited by the researcher for their content, rather than to illustrate some other feature of electronic messages that is largely out of the (conscious) control of the author (e.g., the configuration of mailer headings or the linguistic means of expression), and (2) the researcher is using the message in a way that its author would approve of and wish to be associated with. Similarly, the view that no identifying characteristics of participants should be revealed assumes that (1) researchers are interested in general patterns of participant behavior rather than specific patterns of, for example, language use (and thus that paraphrases are adequate for the purposes of the investigation), and (2) researchers would otherwise "exploit" the self-revelations of participants for personal gain, although it is ethically wrong for them to do so. Clearly the proponents of neither of the two proposals had language scholars in mind (Cavazos is a lawyer, and King a student of psychology); if they had, they would have made different assumptions about what researchers are interested in with respect to computer-mediated messages. Moreover, neither proposal allows for the possibility of legitimate critical research; rather, both assume that there is (or should be) a consensus between investigator and investigated. More problematic yet, both assume a consensus model of interaction among participants, whose needs and interests are represented as essentially homogeneous, albeit different in the two proposals. Such assumptions ignore the insight of social theorists—of whom Michel Foucault is perhaps the best known representative—that discourse is often a site of struggle among groups with conflicting agendas, as well as the empirically observable fact that conflict is a salient characteristic of much CMC. Thus on the level of their underlying assumptions as well, the two proposals do not reflect the complex reality of cyberspace and cyberspace scholarship.

My purpose in the present essay is to problematize these two proposals by demonstrating that they exclude legitimate forms of CMC research. Specifically, I argue for the legitimacy of both linguistic research and critical research in computer-mediated contexts, and focus attention on ethical issues raised by each. I do this by considering each type of research separately (there is no necessary relationship between the two), with theoretical background drawn from various disciplinary practices in the social sciences, especially the field of linguistics, and illustrations drawn from my own research into gendered language use in Listserv discussion groups on the Internet. In concluding, I revisit the practical question of ethical guidelines, and suggest that in the face of the daunting task of formulating a single set of guidelines that would not unfairly discriminate against some legitimate forms of research, existing guidelines within each discipline should be followed, with the addition of a CMC-specific guideline specifying that restricted-access electronic forums have different rights and obligations than do those that permit public access.

Issues Raised by Language-Focused CMC Research

If any group of scholars ought to be interested in CMC, it is linguists. Indeed, CMC is arguably the greatest boon to the study of language use since the invention of the portable tape recorder in the 1950s. Like the tape recorder, it makes possible the analysis of naturally occurring communication on a scale that was previously unimaginable. Before the tape recorder, linguists had to write down speech they had heard from memory, a fact that seriously limited the amount of verbatim material that could be transcribed and analyzed. With the advent of the tape recorder, entire discourses (conversations, stories, speeches,

etc.) could be recorded and transcribed by the analyst at leisure, resulting in larger corpora and enabling discourse-level patterns of usage to emerge that were previously invisible. Thus the field of discourse analysis was born in the 1960s and 1970s (van Dijk, 1988).

However, transcription is tedious and time-consuming, and these practical constraints limit the amount of material that can be analyzed by any one linguist. In contrast, CMC is pre-transcribed—participants have typed in the "data" themselves. And CMC is plentiful, a fact that, in combination with advances in computer-assisted corpus analysis, enables the construction and analysis of much larger and more diverse corpora than were previously possible (see, e.g., Collot & Belmore, 1993; Yates, 1993). Last but not least, CMC is socially situated in "virtual communities" (Rheingold, 1993), the workings of which are rendered explicit as users negotiate new norms of behavior in virtual environments—all of which necessarily takes place through language. Thus computer network technology makes possible more and better (including more socially contextualized) language research than was previously possible.[1]

A comparison between CMC and the tape recorder is instructive with respect to research ethics as well. The availability of tape-recorded data in the early days of discourse analysis research raised ethical questions for language researchers similar to those being debated for CMC today. Could speakers be recorded without their knowledge, i.e., to obtain more natural data? Could transcripts of recorded speech be used in research publication if the speakers were no longer available to be asked for their permission, because they had moved away, for example, or because the recording had been made by someone not personally known to the researcher? If transcripts were quoted verbatim, should speakers be identified by name? In response to these questions, a rough and ready set of conventions has come into use in spoken language research, conventions that arise out of specific ideological commitments characteristic of linguistics and related disciplines.

One such convention concerns the collection of naturalistic spoken language data. The field of linguistics has long been concerned with the "Observer's Paradox," that is, the problem of how to collect authentic data without the collection process interfering with the phenomena observed (especially articulate statements of this problem and some possible solutions were voiced by sociolinguist William Labov as early as 1966). Covert tape-recording has obvious advantages in this respect, since if people are unaware that a researcher is recording their speech, they are less likely to become self-conscious, correct (or overcorrect) in the direction of prescribed norms of usage, or otherwise produce unnatural speech. Ethically, covert tape-recording is considered acceptable in public contexts; a well-known example is Don Zimmerman and Candace West's 1975 study of cross-sex interruptions, in which the researchers recorded conversations between couples overheard in drugstores, coffee shops, airports, and other places to which "any member of the public" could have "normal access." In private contexts, in contrast, ethical considerations dictate that researchers should inform speakers beforehand that they are being recorded, and the tape recorder should ideally be visible. (At least, this is the advice I give students before sending them out to collect data for discourse analysis classes.) However, covert tape recording may be justified even in private contexts, for example, if the knowledge that they are being recorded could make speakers self-conscious to the point of not producing the linguistic phenomena under investigation. A justification of this sort is given by Penelope Harvey (1992) for recording the drunken speech of Quechua-speaking Indians in a small mountain community in Peru; the informal and often irreverent drunken speech would be self-censored in the "official" presence of the tape recorder, although not in the presence of the participant observer, especially if she too were drinking.

Although they have been widely accepted, neither of the research practices used as illustrations in the previous paragraph is entirely ethically unproblematic, especially in light of the considerations raised by King in this issue with respect to CMC research. The conversations overheard by Zimmerman and West were never intended as public; they were private conversations between couples that happened to take place in public settings. And the Peruvian Indians clearly intended their drunken speech to be off-record and ephemeral, as Harvey herself notes. What renders such practices more or less acceptable in spoken language research is that there is a convention of representation (e.g., in writing up the research for publication in a journal), according to which the actual identities of the speakers are disguised. In linguistics research papers, speakers are almost never identified by name as being the source of data presented; pseudonyms may be used, or generic labels such as "a fifteen-year-old boy," "a female associate professor," etc. Most commonly, examples are unattributed, beyond a general description of the data corpus in the methodology section of the paper.[2] Thus any given speaker could plausibly deny that she was the source of any examples used, and no one could prove otherwise.

Admittedly, in some cases the disguise is rather transparent, especially to people with insider knowledge of the speech situation being described. An example of this is found in a 1974 article by Charles Fillmore, (now) professor emeritus of linguistics at the University of California at Berkeley, in which he describes a handwritten message posted on the door of one of a colleague's two offices. As a linguistics graduate student at Berkeley at the time I read the article, I immediately recognized—with some amusement—which faculty member he was referring to, although he had used a pseudonym: the professor's initials. However, the majority of readers presumably would not have access to this information, and thus the identity of the source was masked for most audiences.

One might wonder why, in a situation such as the one just described, the source of the datum was not named. After all, the message was posted on an office door in a public university for any and all to read, and the professor's name was displayed on the door. One reason academic linguists are so little concerned to link sources with their words (aside from professional delicacy, perhaps, when the source is a colleague) is that the focus of linguistic investigation is generally on the form rather than the content of the utterances. Linguistic form is considered something speakers do not consciously produce, as much as it reflects a general competence they possess as speakers of their language. Correspondingly, their individual identities and linguistic quirks are generally of less interest than their membership in the group of speakers whose language is being studied. This is true for the couples recorded by Zimmerman and West (what was most relevant in that study was whether speakers were male or female) and for the Peruvian Indians recorded by Harvey (who were being studied as Spanish–Quechua bilinguals). If the researcher were to give detailed information about individual speakers in writing up such research, it would most likely be perceived as irrelevant and distracting. Similarly in Fillmore's article, although it was amusing to me to recognize a covert reference to someone I knew, the identity of the professor who posted the message on his door was of secondary importance compared with what the linguistic form of the message was intended to illustrate about the English language.

A further practical consequence of focusing on form is that the content of examples quoted in linguistics scholarship is often banal, fragmented, or both. This provides speakers with a certain protection as well. A speaker is unlikely to feel concern at being represented (anonymously and out of context) as having said, "I was there for about uh six . . . six years";[3] for one thing, she can deny it was she who said it, and for another, who would care even if she had? Admittedly, the situation becomes more complicated when

someone is quoted as interrupting or being drunkenly challenging of another, as these are language behaviors that carry social stigma in western middle-class society. Nonetheless, I have never heard of a case where a speaker complained of how she was represented in example sentences in a linguistics publication,[4] nor of harm befalling any individual as a result of such representation. The credit for this is probably due to the fact that in writing up linguistic research, regardless of what has gone before in the research process, the anonymity of speakers is generally preserved, not just in how they are referred to, but also in terms of the content of examples selected for representation.

What, then, of CMC? It is not difficult to imagine computer-mediated situations parallel to those just described for spoken language research. Much CMC, such as that on Usenet newsgroups and on open-subscription Listservs, resembles Zimmerman and West's conversations in public places—researchers can easily "overhear" it, although they may not have been the intended audience, strictly speaking. Harvey's drunken discourse, in contrast, more closely resembles a Listserv group with a restricted membership, where what is said is only intended for the members of the group, although the researcher may be a participant observer in the group, and in that role part of the intended audience. Treating CMC like spoken conversation, one could argue that as long as the anonymity of participants is preserved, it should be ethically possible to cite fragments of electronic messages from virtually any source. However, this view is problematic, precisely because CMC is typed rather than spoken, and leaves a physical record that can be archived or otherwise preserved. As a consequence, it is much more difficult, in practical terms, for a researcher to insure absolute anonymity—a determined reader of the published article, armed with the name of the group, could trace the message and discover the "real" identity (that is to say, the login name) of the e-mail account that originally sent it, for example, by searching the archives, if such are available, for keywords contained in the message. The likelihood that anyone would actually bother to do this may be negligible for examples published, e.g., in an article on spelling conventions in Internet Relay Chat (Werry, in press) or on the use of pronouns in academic computer conferencing (Yates, in press). However, to insure absolute anonymity, we might adopt King's suggestion that computer-mediated groups not be identified by name or any other distinguishing feature; this, in combination with disguising the identity of the message poster, would make it exceedingly difficult for anyone to discover the message source.

But is spoken discourse the best analogy for CMC? CMC is, after all, typed; it can be edited, and it leaves a (potentially) enduring record. An alternative is to treat CMC like written material. In the linguistics literature, examples drawn from published written sources are given full citations, in accordance with copyright law, and are subject to "fair use" requirements. If all CMC is copyrighted, as Cavazos (1994) and others (e.g., Gurak, 1996) have claimed, researchers investigating computer-mediated language should be able to cite any of it they have legitimate access to, as long as they explicitly credit the source.

The crucial question then becomes: Is CMC more like spoken or written discourse? Linguistic research suggests that it is intermediate between the two (Collot & Belmore, 1993; Yates, 1993, in press), and thus that any direct analogy to either speaking or writing is too simplistic. Indeed, my own research practice treats CMC in neither of the two ways just outlined. First, I do not use individual message senders' real names. There are scholarly reasons for this—individual variation is not the focus of my research—as well as ethical reasons—I wish to reserve the right to critique the discourse that I analyze (see later discussion), and in order to do so without damaging the reputations of the individuals who happen to provide me with data, I anonymize them. In this, my practice in citing CMC sources follows that for spoken language research.

At the same time, I identify (open-access) discussion groups by name. I follow this practice (and indeed encourage it in others) for two reasons. First, it strengthens the quality of the scholarship by providing concrete detail, which not only adds informativity (those who are familiar with the group can access and apply their own knowledge of it), but also allows the empirical claims of the work to be independently assessed. That is, the reader needn't take my word for it that there are gender differences of the sort I claim in group X; group X is open to any interested party; in principle, they can subscribe and observe for themselves. Whether or not anyone actually does this is immaterial; what is important for empirical research is that its results be potentially reproduceable by others. If I were to mask the identity of the group, my claims could not be directly evaluated; they would have to be accepted (or not) on the basis of other qualities of the work, such as its rhetorical persuasiveness. The second reason has to do with the type of CMC I analyze, which is primarily that of open-access Listserv discussion groups, many of which have an academic focus (see, e.g., Herring, 1993, 1996, in press a; Herring, Johnson, & DiBenedetto, 1995). This discourse has a flavor that is strongly public, even exhibitionistic at times—it is apparent that many individuals post with an audience in mind, aiming to persuade and impress others with their eloquence and reason. While we might not wish to claim that all messages posted to such groups are "publications," that is, intended to endure through time, it seems entirely appropriate to compare them to public broadcasts, which are designed to reach a wide audience at a particular point in time. (This comparison holds even more strongly for Usenet, where the precise extent and nature of the audience for any given message is unknowable.) As broadcast material, the content of electronic posts is in the public domain, and there is thus no reason not to indicate the group they were broadcast to and through. In this respect, my practice more closely resembles that for written language research.

It may seem that there is an inconsistency in this stance—group names are public information, but individuals' names are not. However, this inconsistency is more apparent than real. In principle, when individuals choose to broadcast their messages to public forums, their names become public information. Researchers may, as a courtesy, avoid using real names in their actual research practice, such as in writing up research that represents participants' messages unflatteringly. However, this is neither a legal nor an ethical requirement. My practice of not mentioning names also fits with a broader ideological preoccupation in linguistics research, namely, that what is important are patterns across groups of speakers, rather than individual linguistic variation. Masking a participant's identity, even if it does not actually "protect" him or her from being recognized by some, is a conventional means of signaling that the identity of the person who posted the message is secondary to the person's membership in a larger social grouping that uses language in characteristic ways.[5]

Astute readers may have noticed a positivist bias pervading this discussion of language research, including in statements of my own scholarly values. Positivism, which Cameron et al. (1992, p. 6) define as "a commitment to the study of the frequency, distribution, and patterning of observable phenomena," is concerned with producing testable claims and procuring value-free observations in a scientific manner. Correspondingly, underlying much linguistic methodology is a fundamental mistrust of the "subjective," which includes speakers' self-reports of their language activity, as well as data "contaminated" by the involvement of the linguist herself (hence the "Observer's Paradox"). Positivist approaches have produced much valuable research, including some that identifies patterns in computer-mediated language and addresses the important question of how CMC compares with other modalities of human communication (see, e.g., Ferrara et al., 1991; Section I in Herring, in press b). Moreover, such research poses little threat to the

well-being of the researched subjects or their communities, since the researchers observe from a distance, preserve the anonymity of subjects, and often focus in rather narrowly on linguistic phenomena such that even when examples are directly quoted, little or no personal information about the sources is revealed.

I submit that such research is legitimate, and that its requirements should be taken into consideration along with those of other research types when discussion of ethical guidelines for CMC research arises. Researchers working within linguistic traditions must be allowed to cite examples verbatim in order to identify and illustrate the phenomena under investigation, contra King's proposal that paraphrases be used instead. At the same time, it is inappropriate to require such researchers to provide full citations for all sources, as Cavazos proposes, just as it would be inappropriate to require linguists in studies of spoken language to identify the individual source of each example sentence. Not only is such information generally irrelevant in research focused on linguistic form, but the requirement could have a chilling effect on linguistic research, in that researchers would tend to avoid research topics that have any potential to make their sources feel self-conscious when the results are published. Moreover, such restrictions are unnecessary, in that research in the tradition of linguistic positivism tends to pose minimal ethical problems as long as subjects are represented anonymously. Problems only arise in that the possibility of such research is not forseen by either the proposed "copyright" or "perceived privacy" guidelines.

Issues Raised by Critical CMC Research

The linguistic research described in the previous section illustrates one model of social science research that is currently being extended to communication in computer-mediated contexts. However, not all social science research assumes that it is possible or desirable to produce "objective" knowledge by maintaining a distance between researcher and researched. Nor do all research paradigms share the view that researched individuals and communities must remain untouched by the research; some allow for active researcher intervention on behalf of the researched population, or for giving the researched a say in the (co-)construction of the research itself.

This latter view, which is grounded in the philosophical precepts of relativism (Cameron et al., 1992), is argued for with respect to CMC research by Christina Allen in this issue. Relativism does not privilege the practices and beliefs of one group over another, or of the researcher over the researched, but holds all to be equally valid (Cameron et al., 1992). Accordingly, Allen calls for increased interaction between researcher and researched in the CMC research process, not just in obtaining advance informed consent, but in letting the researched speak for themselves through interviews, and in giving them opportunities to "correct" or change what the researcher is writing about them before it goes to press. However, this set of recommendations, too, is highly problematic if generalized broadly. Specifically, like the other proposed guidelines discussed thus far, it renders critical analysis difficult or impossible, in that it presupposes a consensus view of the researcher–researched relationship and of CMC more generally.

The consensus view flies in the face of the experience of many users that CMC is riddled with conflict. Groups conflict with groups (misogynists with feminists, white supremacists with liberals, expatriate Turks with expatriate Armenians, personal users with commercial advertisers, civil libertarians with advocates of regulation, Chomskyan linguists with functional linguists, etc.), and individuals regularly enter into conflict with other individuals on Usenet newsgroups, chat channels, and academic

Listservs alike. Various explanations have been proposed for the high incidence of conflict in on-line interaction, ranging from "disinhibition" caused by the depersonalizing nature of the medium (e.g., Kiesler et al., 1985; Kim & Raja, 1991), to a positive valuing of conflict as a form of (gendered) social interaction (Herring, 1994, 1996). Whatever its explanation, the prevalence of conflict has as a consequence that users, even those subscribed to special-interest discussion groups, cannot reasonably be considered homogeneous populations with respect to their interests and social/political agendas.

The heterogeneity of CMC raises problems for many of the ethical recommendations proposed in this issue. For example, Waskul and Douglass in their article recommend that CMC researchers obtain informed consent and work only with key informants. But, informed consent from whom? To get all participants to consent to any project, no matter how unintrusive, is a difficult task. If the project is at all controversial, the chances that everyone will agree are virtually nil. Should the researcher then abandon the project? An alternative that is sometimes proposed is to obtain the consent of the listowner on behalf of the group as a whole. However, this does not "give the researched a say"; rather, it privileges the perspective of a single individual who occupies a position of power within the group. Finally, the notion of "key informants" is problematic as well. Allen in her dissertation research on LambdaMOO ended up working closely with only 4 out of 9000 group members. How representative were the views of those four individuals of that complex community as a whole? Methodological choices of this sort essentially limit the kinds of research that can be done to case studies, valid in and of themselves, but surely not the only kind of CMC research worth doing.

Suppose for the sake of argument, however, that permission is granted by group consensus for a researcher to observe and analyze the discourse of a "virtual community" (an expression that is itself ideologically loaded, in that it assumes a shared group identity). Is the researcher then obliged to insure that whatever he or she writes meets with the approval of every member of the group, or is at least inoffensive to them? What if the project reveals a political division within the group, or patterns of dominance of some members by others? Under such circumstances, the research findings, if honestly represented, will likely make some participants uncomfortable. Should results of this sort then not be written up, or be represented less than honestly? Such suggestions are clearly unpalatable on scholarly grounds.

In short, guidelines that require consensus are both unrealistic and unreasonably restrictive. What ethical guidelines, then, should be followed by researchers who investigate relations of power and dominance in cyberspace? Such critical research is important and necessary if scholarship is to involve itself in the "real" (in the case of CMC, virtual) world, rather than insulating itself in academic abstraction. However, any involvement with the "real world," with its real people and real social concerns, is likely to be ethically complex.

As with positivist scholarship, the ethical concerns that arise in critical scholarship are linked to the research paradigm that provides the methods, goals, and theoretical assumptions for the research. One such paradigm that I have followed to some extent in my own research is what I will call social realism, modifying a term from Cameron et al. (1992). Social realism holds that different social groups, as defined, for example, by gender, race, and class, are characterized by an unequal distribution of power, such that some groups dominate and others are dominated by them. Viewed from this perspective, the heterogeneous, often conflictual nature of computer-mediated interaction is a reflection of multiple interests and struggles for power.

Power, in turn, is negotiated primarily through discourse, especially in the kinds of "official" discourse that create what comes to be defined as knowledge in a culture (Foucault, 1980). Academic researchers participate in creating official knowledge, especially when the researched are members of less powerful groups such as crime victims, the mentally ill, children, homosexuals, ethnic minorities, etc., and the research contributes to labeling the behaviors of the researched in various ways (Cameron et al., 1992). In research in this paradigm, it is naive to claim that we can be "objective" or "neutral"; researchers as well as researched subjects are socially situated actors with their own personal and political agendas.

Some scholars, including many feminists (for example, Steiner, 1989), have responded to the ethical challenges raised by social realist research by directly acknowledging their own biases and potential biases when presenting their research. This may include, if one is a feminist, explicitly acknowledging an activist agenda to "critique and to eliminate women's oppression and the oppression of others" (Steiner, 1989, p. 158). It may also include incorporating practices into the research process that are designed to empower or otherwise benefit the researched group. Thus Deborah Cameron (1992) describes a project in which she worked with Afro-Caribbeans in a London youth club to produce an anti-stereotypical video about racist language. On the basis of this experience, she proposes a number of guidelines for carrying out "empowering" research; these include soliciting the views of the researched group about the phenomena under investigation, sharing knowledge and research tools with the researched, and presenting the results of the research in a way that the researched will find accessible (p. 128).

However, Cameron's guidelines assume that one is working with less powerful groups, and that one wishes to actively benefit those groups. But what if the researched group is doing harm to others? In such circumstances, the position of the researcher may well be critical rather than supportive. A recent example of a study of this sort is that of Peter Adams, Alison Towns, and Nicola Gavey (1995). Male researchers orally interviewed men who had previously been arrested for violence against women concerning their attitudes about male dominance. The resultant journal article displays a number of features of social realist research. The authors indicate their personal subjective interest [the project was triggered by an incident within their profession (clinical psychology) that disturbed them], and they explicitly acknowledge having a profeminist theoretical stance. However, they do not seek to advance the agendas of the interviewed men. On the contrary, although they indicate that they listened politely to the subjects, they deemed it necessary, at the end of each interview, to explicitly question attitudes that supported violence against women, lest the interview process itself be seen to encourage further violent behavior. Similarly, Harvey (1992) notes the importance of directing empowering research "as much at the political consciousness of the powerful as at the powerless."

Many of the same considerations arise in the research I have conducted on gender patterns in CMC. The high incidence of conflict in cyberspace makes it an ideal setting in which to analyze the discursive construction of power, and to seek answers to such questions as "whose interests are worthy of debate, who gets to talk, and who is regarded as an effective communicator to whom others must listen?" (Steiner, 1989, p. 158). Moreover, the answers to these questions have important real-world implications, in that they potentially limit access by some groups to computer-mediated information and interaction. Thus I set out 5 years ago to find an answer to the question of why women tend to participate less than men on mixed-sex academic Listserv discussion groups.

I hasten to add that I did not start out with the intention of conducting critical research; my training and background were that of a positivist linguist, not a social theorist.

I did, however, believe that empirical methods could be fruitfully applied to analyzing situations of broad social importance, and thus in order to address the question of gender imbalance in Listserv participation, I subjected electronic interactions to linguistic text analysis, sent out questionnaires, and conducted extensive ethnographic observation of Listserv groups. In the process, I was confronted with a number of problems that are typical of social realist research, and that effectively transformed my research practice. I identify five of these here.

The first problem was that of my own subjectivity. I began by analyzing a discourse community with which I was closely professionally involved, a Listserv discussion group for academic linguists. Moreover, my initial research question was one in which I had a personal stake as a female academic. Although I have made every effort to be rigorous in my methods and interpretations, I obviously can not claim a "neutral" stance toward the topic of investigation. Inevitably, my results reflect the perspective of a female academic linguist computer network user.

The second problem I faced was whether or not to obtain informed consent from the group I studied. I did not ask the group's permission to observe or analyze their interaction, as it seemed clearly to be public discourse; anyone could join the group and read the messages posted to it by sending a "subscribe" message to the Listserv, and the topics of discussion were generally impersonal and academic. Moreover, I suspected I would not get the approval of all involved, and I did not want to be constrained to write only what (the more vocal, dominant) members of the group would approve. For what I had discovered—reinforced by the comments of respondents to the questionnaires—was that lengthy and often tendentious messages posted by a minority of male subscribers effectively set the terms of the discourse for the group as a whole, and intimidated others—especially women—into silence.

It was apparent to me that some of these male subscribers were doing considerable harm, albeit probably unintentionally for the most part. In the interests of raising the general consciousness about gender inequality on the Internet, I wanted to expose the behavior of these men to public scrutiny. Thus the third problem became one of representation—should I use the real names of subscribers who posted messages to the group? In the end, I decided to accord all participants, male and female, the same degree of anonymity, as my criticism was not targetted against individuals but rather against gendered patterns of social interaction. I did, however, identify the groups by their real names, to allow others to be able to verify my claims.

The fourth problem arose when it came time to publish my research results. Clearly, it was in my best professional interest to publish in academic (print) venues recognized by my discipline and valued in my university's tenure and promotion procedures. However, soon after I began publicly presenting my results, it became apparent that they had real-world consequences. Many women reported feeling empowered by the research; it validated experiences they had had on-line but could not previously name, and they took the results and discussed them in other electronic groups, using them in some cases to draw attention to and resist discursive domination (as I was later to learn). These effects introduced a new ethical consideration: I began to feel increasingly responsible to make the results acessible to others who might benefit from them, including academic computer users outside my discipline as well as nonacademic users and potential users. Accordingly, I have made some of my results publicly available on the Internet, and have spoken about my research to a variety of audiences, including librarians, computer scientists, and public school teachers.

Fifth and last, I encountered a problem that is inherent in conducting research on any socially subordinate group (such as women on the Internet), namely, that the results of

one's research may be used by others to perpetuate negative stereotypes about the group. Early on, I was aware of the potential for my research results to be used to trivialize or marginalize women in cyberspace. For example, the finding that many women are intimidated by tendentious male posting behavior (Herring, 1993, 1995) could be taken to support sexist views that women are emotionally fragile and unfit for life on the virtual frontier. (As one man posted in response to a complaint, "If you can't take the heat, ladies, get out of the kitchen."). The finding that women's electronic messages tend to be more polite than those of men (Herring, 1994, 1996) could be taken as evidence that women are indirect or insincere (i.e., they seem to agree when they really disagree, unlike men who express honest disagreement), and the finding that women's but not men's messages tend to exhibit social alignment (Herring, in press) could be distorted to support the view that women only use computer networks for "socializing," while men exchange "information" (e.g., Kantrowitz, 1994). It is difficult to control how one's work is used once it enters the public domain—just as it is difficult to control the fate of one's electronic messages once they are publicly posted. However, I have tried to avoid facile generalizations in my professional speaking and writing that could contribute to the popular stereotypes about men, women, and computers that have begun to surface in the mainstream media and in cyberspace itself. This is an ongoing challenge.

The research just described could not have been carried out had I been obliged to follow strictly positivist or consensus-based guidelines. The subjectivity of my personal situation vis-à-vis my research topic meant that I could not do "objective" scholarship, which for many positivists means no scholarship at all. Had I decided to undertake the project regardless, I would most likely have been stopped short by the requirement of informed consent—individuals engaged in socially stigmatized behavior would have been unlikely to agree to a critical examination of that behavior. Finally, if I had carried out the research following the injunction not to include any particulars regarding the source of my data, the entire project and its conclusions could have been invalidated on the grounds that they were not verifiable. Ironically, using individual's real names rather than pseudonyms could have a similar effect—behaviors can more readily be viewed (and dismissed) as characteristic of only a few individuals, rather than as gender-related, when recognizable names are used.

Nor could those same positivist and consensus-based guidelines have prepared me for the ethical responsibility I felt when my research results turned out to have social consequences. This responsibility can be characterized as a concern, not just to avoid harm, but to do good, for the researched population and for society as a whole. Most research in the social sciences is motivated, fundamentally, by a desire to find solutions to problems (such as crime, poverty, racism, sexual abuse, inequality) and to contribute to social change. Similarly, inasmuch as male dominance undemocratically restricts the participation of women in cyberspace, it is a social problem to be solved, and research can potentially play a useful role. Given this, we may ask whether a lack of concern with the larger social effects of one's research is an ethical failure. Would it have been inethical (or merely elitist) of me to address my research exclusively to academic audiences? My professional credentials might be more impressive had I done so. Would it be inethical (or merely self-centered) of me to "sell out" by writing popular books about men and women on the Internet that fit current cultural stereotypes? I might be wealthier if I did so. But women on the Internet would benefit less (and might experience harm, e.g., from negative stereotyping); thus these must be considered ethical issues. Yet the ethical guidelines proposed for CMC research by King and others do not address them.

If computer-mediated communication is to be free and democratic, cyberspace cannot be dominated by groups that restrict the access of others. Nevertheless, the existence of multiple agendas and struggle in cyberspace means that the potential for such dominance is always present. In addition, the Internet will likely face the threat of external dominance, including by economic interests, in the future. Researchers have a responsibility to critique these aspects of CMC, as they have leveled criticism at other forms of mass communication in the past. Indeed, failure to do so could reinforce the hegemony of dominant groups, since theirs are the voices most likely to determine the appearance that any "consensus" will take. Thus the consensus-based guidelines of King, Waskul and Douglass, and Allen, inasmuch as they inhibit critical research practices, are themselves ethically problematic.

Summary and Conclusion

I began this article by comparing two sets of proposals regarding ethical conduct in CMC research, one based on the notion that CMC is "published" material, and the other on the notion that CMC is "private" interaction. I pointed out that these proposals, although they make contradictory assumptions about the nature of CMC, make similar assumptions about what constitutes research: namely, that it focuses on content rather than on form, and that it is (ideally) consensual. In the previous two sections, I argued that these assumptions are inappropriately narrow; specifically, they exclude research practices in the linguistic and the critical traditions (and no doubt other traditions as well), each of which raises different ethical and scholarly considerations. These considerations are best addressed within each academic discipline.

In this section, I return to consider (briefly) the nature of CMC. It is by now a truism that cyberspace is vast and diverse. The obvious answer to the question of whether multi-participant CMC is more like published text or ephemeral private communication is that it is both, at different times and in different places, and other things besides—it is also soapbox rhetoric, cocktail party conversation, idle chat around the copy machine, a stag party, group therapy, playing Dungeons and Dragons, attending an academic conference, etc. Thus just as no single set of disciplinary guidelines is appropriate for all research paradigms, it is difficult to imagine any single set of guidelines that could appropriately reflect the nature of the interaction in all of these different genres.

At the same time, however, the use of CMC as material for scholarly research gives rise to a unique ethical concern. The "on-record" nature of electronic communication and the ease with which it can be redistributed mean that the privacy of researched subjects, in situations where privacy is a legitimate concern, is more difficult to ensure. This suggests that some CMC-specific guidelines may be necessary.

Fortunately, there is a simple means to address ethical issues associated with the public/private debate. That is to recognize the de facto public nature of most multi-site CMC (based on the fact that anyone with access to an e-mail account and other commonly available software can subscribe to Listserv groups, read Usenet messages, join chat channels and participate in MUDs and MOOs), and openly declare such varieties public as the default. Conversely, private arrangements must be explicitly set up and managed; on the Internet, these might include Listserv groups that require the approval of a listowner to join, invitation-only chat channels, and "private rooms" in MUDs and MOOs.

The ethical prescription would then be straightforward, and it would apply not only to researchers, but to any individuals wishing to observe and report on electronic interac-

tion: Public interaction is repeatable for any reasonable and nonmalicious purpose (with citations of the source where credit for ideas is due), but private interaction should not be repeated outside the group without explicit permission from the sources involved. An advantage of this system is that the technology is already available, such that any group that so wishes may restrict access and designate itself as private (it is important to do both—simply declaring one's group "private" without any means of controlling access is not likely to be credited). Thus there would be no excuse for not doing so, once the distinction becomes sufficiently conventionalized.[6] In the meantime, considerate researchers will avoid exposing to the public gaze interaction on sensitive topics such as that of the sexual abuse survivors described by King. Ultimately, however, it will be the responsibility of such groups to convert to a private format (or else decide to live with the risk of exposure if the publicness of their existing arrangement is advantageous to them for other reasons).

By way of analogy, I note that when individuals broadcast by appearing on television or talking on the radio, they cannot see their audience, nor do they know everyone who is in that audience. Indeed, they may only be interacting with a small number of people (e.g., a talk show host and other guests, perhaps in the presence of a small studio audience), which might in principle give them the illusion of carrying on a private or contained conversation. King and Waskul and Douglass advance this kind of argument to claim that researchers should respect the "perceived privacy" of individuals posting messages on open-access Usenet and Listserv groups. However, it would be ludicrous to make such an argument for television or radio broadcasts, because their public nature and the conventions of broadcasting are understood by all normal, educated adults in industrialized societies. Such a widespread understanding of the public nature of CMC has not yet been achieved, but it could very well be achieved in the future, given increased exposure to CMC by greater numbers of people, and modest efforts to educate users regarding the default public nature of computer-mediated communication.

Underlying this proposal is an assumption that participants and listowners have a responsibility to themselves to protect their own privacy. Researchers share in this responsibility, of course, given their more powerful position in the relationship, but researchers have other responsibilities as well, to themselves and to their research. To expect researchers to be solely responsible for the interests of the researched invites paternalism and does not guarantee that the best interests of the researched will necessarily be served.

In conclusion, researchers should be actively concerned to avoid causing harm to researched populations as a result of their research, in cyberspace as elsewhere. However, ethical guidelines and conventions of practice already exist within individual research traditions that are largely adequate for this task. Moreover, a means for protecting the privacy of researched subjects in computer-mediated contexts is already available and can be developed further, such as by exploiting the technical distinction between open-access and restricted-access groups. It is unnecessary to implement invariant guidelines that severely limit the range of legitimate research practices to those with which some subset of scholars are most personally familiar. Cyberspace is a complex phenomenon, and to understand it fully will require a diverse set of research practices. This essay thus constitutes a call to balance ethical considerations with a broader conceptualization of CMC research, by recognizing its current diversity and its potential to contribute further to our understanding of computer-mediated interaction in the future.

Notes

1. Unfortunately, not many linguists have realized this yet. But then, it took more than 10 years after the invention of the tape recorder for the implications of that technology to begin to be realized in a new field of study—discourse analysis—so there is hope that the enormous potential of computer-mediated language analysis will yet be realized.

2. Two exceptions to the non-use of source names are (1) when the source is a well-known public figure (e.g., Rush Limbaugh), and (2) when the researcher has asked and obtained permission from the sources to use their real names (usually in situations where there is a small number of sources, and where the researcher wishes to give a more personal flavor to the research; however it is not considered necessary or even especially desirable to do this). A third situation in which real names may be used is when language consultants and research assistants who have provided data are named and thanked in a footnote; this is especially common in analyses of lesser studied languages.

3. This example is given in Chafe (1994, p. 205).

4. This does not of course rule out complaints about the correctness/appropriateness of the examples in terms of the linguistic analysis, or complaints about how groups of people (such as women, homosexuals, ethnic minorities, etc.) are represented in linguistic example sentences more generally.

5. I understand Fillmore's motives (in his 1974 paper) to be similar. He could very well have named the professor who posted (broadcasted) the message on his office door; indeed, later in the same article he names another such "real-life" source. However, the individual identity of the first source was not crucial to the linguistic point he was trying to make, and the example additionally made the professor look vaguely ridiculous—hence the superficial disguise.

6. Conventionalization is the key to the success of such a proposal in ensuring the privacy of those who legitimately claim it. It may ultimately not be practical or possible to enforce absolute privacy in any computer-mediated environment; any time a message is sent to an unseen recipient, the potential is there for its content to "leak" to other audiences, accidentally or maliciously, no matter what technical or social means of enforcement are employed. However, most people who wish to be considered respectable by their peers live according to the social contract, and thus would tend to honor clearly signaled conventions of privacy in a computer-mediated context. To further encourage compliance, malicious failure to do so would need to carry professional and social stigma, and perhaps be legally actionable as well.

References

Adams, P., Towns, A., and Gavey, N. 1995. Dominance and entitlement: The rhetoric men use to discuss their violence towards women. *Discourse and Society* 6(3):387–406.

Cameron, D. 1992. "Respect, please!": Investigating race, power, and language. In *Researching language: Issues of power and method,* eds. D. Cameron, E. Frazer, P. Harvey, M. B. H. Rampton, and K. Richardson, pp. 113–130. London: Routledge.

Cameron, D., Frazer, E., Harvey, P., Rampton, M. B. H., and Richardson, K., eds. 1992. *Researching language: Issues of power and method.* London: Routledge.

Cavazos, E. A. 1994. Intellectual property in cyberspace: Copyright law in a new world. In *Cyberspace and the law: Your rights and duties in the on-line world,* eds. E. Cavazos and Gavino Morin, pp. 44–66. Cambridge, MA: MIT Press.

Chafe, W. L. 1994. *Discourse, consciousness, and time.* Chicago: University of Chicago Press.

Collot, M., and Belmore, N. 1993. Electronic language: A new variety of English. In *English language corpora: Design, analysis and exploitation,* eds. J. Aarts, P. de Haan and N. Oostdijk, pp. 41–56. Atlanta, GA: Rodopi.

Ferrara, K., Brunner, H., and Whittemore, G. 1991. Interactive written discourse as an emergent register. *Written Commun.* 8(1):8–34.

Fillmore, C. 1974. *Pragmatics and the decription of discourse.* Berkeley Studies in Syntax and Semantics 1. Berkeley: University of California Berkeley Department of Linguistics.

Foucault, M. 1980. *Power/knowledge: Selected interviews and other writings, 1972–77,* ed. C. Gordon. Brighton: Harvester.

Gurak, L. J. 1996. The multi-faceted and novel nature of using cyber-texts as research data. In *Computer networking and scholarship in the 21st century university,* eds. T. M. Harrison and T. D. Stephen. Albany: SUNY Press.

Harvey, P. 1992. Bilingualism in the Peruvian Andes. In *Researching language: Issues of power and method,* eds. D. Cameron, E. Frazer, P. Harvey, M. B. H. Rampton, and K. Richardson, pp. 65–89. London: Routledge.

Herring, S. 1993. Gender and democracy in computer-mediated communication. *Electronic J. Commun.* 3(2), special issue on computer-mediated communication, ed. T. Benson. Available from comserve@rpitsvm.bitnet. Reprinted in R. Kling, ed. (1996). *Computerization and controversy,* 2nd ed. New York: Academic Press.

Herring, S. 1994. Politeness in computer culture: Why women thank and men flame. In *Cultural Performances: Proceedings of the Third Berkeley Women and Language Conference,* eds. M. Bucholtz, A. Liang, L. Sutton, and C. Hines, pp. 278–294. Berkeley: Berkeley Women and Language Group.

Herring, S. 1995. Freedom of speech or freedom of harassment? *The College* (magazine published by the University of Texas at Arlington College of Liberal Arts) 1(1):8–9.

Herring, S. 1996. Posting in a different voice: Gender and ethics in computer-mediated communication. In *Philosophical perspectives on computer-mediated communication,* ed. C. Ess, pp. 115–145. Albany: SUNY Press.

Herring, S. In press a. Two variants of an electronic message schema. In *Computer-mediated communication: Linguistic, social and cross-cultural perspectives,* ed. S. Herring, pp. 81–105. Amsterdam: John Benjamins.

Herring, S., ed. In press b. *Computer-mediated communication: Linguistic, social and cross-cultural perspectives.* Amsterdam: John Benjamins.

Herring, S., Johnson, D., and DiBenedetto, T. 1995. "This discussion is going too far!" Male resistance to female participation on the Internet. In *Gender articulated: Language and the socially constructed self,* eds. M. Bucholtz and K. Hall, pp. 67–96. New York: Routledge.

Kantrowitz, B. 1994. Men, women and computers. *Newsweek* 16 May:48–55.

Kiesler, S., Zubrow, D., Moses, A. M., and Geller, V. 1985. Affect in computer mediated communication. *Hum. Comput. Interact.* 1:77–104.

Kim, M.-S., and Raja, N. S. 1990. Verbal Aggression and Self-Disclosure on Computer Bulletin Boards. Washington, DC: ERIC Clearinghouse on Languages and Linguistics, document ED334620.

Labov, W. 1966. *The social stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.

Rheingold, H. 1993. *The virtual community: homesteading on the electronic frontier.* Reading, MA: Addison-Wesley.

Steiner, L. 1989. Feminist theorizing and communication ethics. *Communication* 12:157–173.

van Dijk, T. 1988. *News as discourse.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Werry, C. In press. Linguistic and interactional features of Internet Relay Chat. In *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives,* ed. S. Herring, pp. 47–63. Amsterdam: John Benjamins.

Yates, S. J. 1993. The Textuality of Computer-Mediated Communication: Speech, Writing and Genre in CMC Discourse. Unpublished PhD dissertation, Open University, Milton Keynes, UK.

Yates, S. In press. Oral and written linguistic aspects of computer conferencing: A corpus based study. In *Computer-mediated communication: Linguistic, social and cross-cultural perspectives,* ed. S. Herring, pp. 29–46. Amsterdam: John Benjamins.

Zimmerman, D. H., and West, C. 1975. Sex roles, interruptions, and silences in conversations. In *Language and sex: difference and dominance,* eds. B. Thorne and N. Henley, pp. 105–129. Rowley, MA: Newbury House.