

**\*This version of the article is a late working copy. Please do not quote exactly without checking with the final published version.**

Critical Analysis of Language Use in Computer-Mediated Contexts:  
Some Ethical and Scholarly Considerations

Susan Herring  
Program in Linguistics  
University of Texas  
Arlington, TX 76019

## THE PROBLEM

In the early years of CMC research (which is to say up until very recently), those of us researching CMC 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 power and exhilaration, but at the same time 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 e-mail messages in scholarly publications, how to cite electronic journal publications on one's curriculum vitae, and other such practical matters. And I especially looked forward to the day when sound advice would become available as to whether (and 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. The decision I made in my first CMC publications was to use pseudonyms or avoid mention of names rather than revealing the actual identities of my data sources, in part because the theoretical orientation of my research is critical of the language patterns used by some participants. My intention has been 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 essay in this volume. 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 intended as 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 (e.g. open debate of academic issues vs. self-revelation of sensitive information in self-help groups) 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 e.g. language use (and thus that paraphrases are adequate for the purposes of the investigation), and 2) the interests of researchers and participants are potentially in conflict, with researchers motivated to "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 power relations are contested in and through discourse, as well as the readily observable fact that conflict is a salient

characteristic of much CMC. Thus on the level of these assumptions as well, the two proposals do not reflect the complex reality of cyberspace and cyberspace scholarship.

My purpose in this essay is to argue for the legitimacy of both language study and critical study in computer-mediated contexts, and to fill a gap in the discourse about CMC ethics by focusing consideration on ethical issues raised specifically by each. I do this by considering each 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.

## 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 1950's. 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 which 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 1960's and 1970's (van Dijk 1985).

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 the "data" in themselves. And CMC is plentiful, a fact which, 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 1992; Yates 1993). Last but not least, CMC is socially situated in "virtual communities" (Rheingold 1993), the workings of which are rendered explicit through users negotiating 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. To be sure, 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.

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, e.g. because they had moved away, 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 which arise out of specific ideological commitments characteristic of linguistics and related disciplines. I briefly discuss two of the most important of these conventions below.

The first 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 also be justified in private contexts, e.g. if the knowledge that they are being recorded could make speakers self-conscious to the point of not producing the linguistic phenomena under investigation at all. 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 described in the previous paragraph is entirely unproblematic, especially in light of the considerations raised by King in this volume with respect to CMC research. The conversations overheard by Zimmerman and West were never intended as public; they were private conversations between couples which 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 identity of the speakers is disguised. Speakers are almost never identified by name in research papers 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., or (most commonly) the example may be unattributed, beyond a general description of the data corpus in the methodology section of the paper.<sup>1</sup> 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 from my personal experience is a 1974 article by Charles Fillmore, (now) professor emeritus of linguistics at the University of California at Berkeley, in which he describes a hand-written message posted on the door of one of a colleague's

two offices. As a graduate student in linguistics 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 actual identity of the source could not have been revealed. 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 also displayed on that door (it is clear from the context given in Fillmore's example that the message was written by the professor). One reason academic linguists are so little concerned to link sources with their words (aside from professional delicacy, perhaps, when a colleague is involved) 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, so 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 minor 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";<sup>2</sup> for one thing, she can deny it was she who said it, and for another, who would care if she had? Admittedly, the situation becomes more complicated when someone is quoted as interrupting or being drunkenly challenging of another; these are language behaviors that carry social stigma, at least 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,<sup>3</sup> nor of harm befalling any individual as a result of such representation. The credit for this is probably due to the fact that in representation, regardless of what has gone before in the research process, the anonymity of speakers is invariably 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 described above 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 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 possible, ethically, 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 which can be archived or

otherwise preserved. As a consequence, it is much more difficult, practically speaking, 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, e.g. 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 an Internet Relay Chat session (Werry Forthcoming) or on the use of pronouns in academic computer conferencing (Yates Forthcoming), but to insure absolute anonymity, we could 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 and others (e.g. Gurak 1995) 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 and Belmore 1993; Yates 1993, Forthcoming), and thus that any direct analogy may be too simplistic. Indeed, my own research practice treats CMC in neither of the two ways outlined above. First, I do not use individual message senders' real names, both for scholarly reasons -- variation at the level of the individual has not been the focus of my research -- and ethical reasons -- I wish to reserve the right to critique the discourse that I analyze (see below), and in order to do so without harming the individuals who happen to provide me with examples, I anonymize them. In this, my practice in citing CMC sources follows that for spoken language research.

At the same time, I identify (public-access) groups by name, thereby enabling readers, should they be so inclined, to access the archives for the groups themselves and identify the real e-mail addresses of individual contributors cited in my examples (although to my knowledge no one has ever tried to do this). 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, e.g. 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). This discourse has a flavor which is strongly public, even exhibitionistic at times -- many people post as though 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 extent and nature of the audience 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. In fact, individuals' names are also public information, to the extent that individuals choose to broadcast their messages to public fora. It is a matter of courtesy, not an ethical or legal requirement, that their real names not be used in research that represents their messages unflatteringly. Thus for ethical purposes, group names and individual names have the same (public) status. The practice of not mentioning names also fits with a broader ideological preoccupation in linguistics research, 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 way of signaling, "the identity of the person who posted this message is secondary to his or her membership in a larger social grouping which uses language in characteristic ways".<sup>4</sup>

Astute readers may have noticed a general positivist bias pervading this discussion of language research, including in statements of my own scholarly values. Positivism, which Cameron et al. (1992:6) 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 assumptions have produced much valuable research, including some which 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; Herring (ed.), Forthcoming). Moreover, such research poses relatively 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 is revealed about the sources.

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 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 the variety of linguistic research that is carried out, in that researchers would tend to avoid research topics that have any potential at all to make their sources feel self-conscious when the results are published. All in all, research in the tradition of linguistic positivism appears to pose minimal ethical problems as long as subjects are represented anonymously. It only becomes problematic in that the possibility of such research is not foreseen 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 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 (Cameron et al. 1992). This latter view is argued for with respect to CMC research by Christina Allen in this volume. 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 problematic if generalized broadly. Like the other proposed guidelines discussed thus far, it presupposes a consensus view of the researcher-researched relationship and of CMC more generally which is incompatible in key respects with the goals of critical analysis.

One of the most striking characteristics of CMC, in my experience, is the extent to which it is a locus of conflict. Groups conflict with groups (misogynists with feminists, white supremacists with liberals, expatriot Turks with expatriot Afghanis, personal users with commercial advertisers, civil libertarians with advocates of regulation, 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 this phenomenon, ranging from 'disinhibition' caused by the depersonalizing nature of the medium (e.g. Kiesler et al. 1985; Kim and 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.

This raises problems for many of the ethical recommendations proposed in this volume. 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? Getting all participants in an electronic forum to consent to any research 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 his or her project? Some of the most interesting research questions that can be asked about CMC involve areas of controversy.<sup>5</sup> The notion of 'key informants' is problematic as well. Allen in her dissertation research on LambdaMOO ended up working closely with only four out of 9,000 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 questions that can be addressed 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". 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 members of the group 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, requiring cooperation and collaboration between researchers and researched in computer-mediated settings is problematic. However, this does not invalidate the possibility of adopting a non-positivist approach in CMC research. In what follows, I discuss an alternative approach, philosophically grounded in social realism, which most closely approximates my own CMC research practice.

Social realism holds that different social groups, as defined e.g. by gender, race, and class, are characterized by an unequal distribution of power, such that some groups dominate and others are dominated by them. Power 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). As researchers, we participate in creating knowledge of this type, especially when the researched are members of less powerful groups such as crime victims, the mentally ill, children, homosexuals, ethnic minorities, etc., and our research contributes to labeling their behaviors in various ways (Cameron et al. 1992). In research of this sort, 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 posed by social realism by directly acknowledging their biases and potential biases when presenting their research. This may include, if one is a feminist, acknowledging an activist agenda to "critique and to eliminate women's oppression and the oppression of others" (Steiner 1989:158). It may also include incorporating practices into the research process which 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 (128).

Somewhat different ethical considerations are necessarily raised by research on members of oppressive groups. A recent example of a study of this sort carried out within a social realist framework is that of Peter Adams, Alison Towns, and Nicola Gavey (Adams et al. 1995). (Male) researchers orally interviewed men who had been arrested for violence against women concerning their attitudes about male dominance. The authors indicate their personal interest (the project was triggered by an incident within their profession (clinical psychology) which disturbed them), and they explicitly acknowledge having a pro-feminist theoretical stance. Moreover, although they do not mention in their article how the results of the research will be used to benefit women, they note that they deemed it necessary, at the end of each interview, to explicitly question attitudes which supported violence against women, lest the interview process itself be seen to encourage further violent behavior. Similarly, Cameron (1992:120, quoting Harvey 1992) underscores the importance of directing empowering research "as much at the political consciousness of the powerful as at the powerless".

There are numerous ways in which CMC warrants research in a social realist paradigm. 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 1989:158). 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. One issue which has attracted considerable attention in the popular media is sexual harassment online (see e.g. Van Gelder 1990; Dibbell 1993); the evidence is mounting that cyberspace is no less sexist than the "real world". There is a need for such issues to be treated seriously and responsibly by social scientists.

My own research on gender patterns in computer-mediated discourse illustrates some of the ethical issues associated with social realism, although I did not start out with the idea of undertaking social realist research. Quite on the contrary, prior to doing my first study, I had no training in social realism, had never read Foucault, had not previously worked on gender issues, and did not call myself a feminist. I believed, however, that positivist methods of linguistic analysis could be used to address issues of wider social importance, and thus I set out to "solve" a problem that was troubling me, namely: why did women participate so little (and so differently) from men on mixed-sex academic Listserv discussion lists? As a female academician who also subscribed to such lists, I obviously could not claim a "neutral" stance with respect to the topic. I did, however, employ empirical methods, including electronic questionnaires and quantitative text analysis of computer-mediated discussions. I also added generous doses of qualitative interpretation (the data do not speak for themselves), drawing heavily on the comments made by anonymous respondents (both female and male) to my questionnaire studies, each of which included open-ended questions. Thus, I solicited and incorporated the views of the researched group, as Allen and Cameron recommend.

However, in other respects, my practice deviated significantly from that for consensus-based research. I did not ask the group's permission in advance to observe or analyze their interactions, nor did I consult with members regarding what I was planning to write. With regard to the first choice, it did not seem necessary to obtain permission to analyze what was self-evidently 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 exclusively academic. As for the second choice, I strongly 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 the questionnaire respondents -- was that lengthy and often adversarial messages posted by a minority of male subscribers effectively set the terms of the discourse for the group as a whole, and intimidated others into silence. When I began presenting this research publicly, it provoked strikingly different reactions from women and men. Women reported feeling empowered by it; it validated experiences they had had online 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 subvert male discursive domination as it was taking place (as I was later to find out). Men, in contrast, tended to respond to my research as though intimidated -- they would remain conspicuously silent while women asked questions and made comments after my conference presentations.

The nature of the response engendered by this research has shaped my research practice in subtle ways, bringing it increasingly in line with social realist practice. I did not start out with the intention of empowering women in cyberspace, but when it became apparent that the work was having that effect, I began to feel increasingly responsible to make it accessible. I have distributed some of the work on the Internet through publication in electronic journals and ftp sites, even though such publications are viewed as less statusful by university tenure and promotion committees. I announce the availability of my papers to the electronic groups I have studied, following up by sending copies (including prepublication copies) to any and all who request them. It has also seemed important to address audiences outside my academic discipline -- librarians, philosophers, computer scientists, public school teachers -- and to participate in formative discourses about CMC such as the present debate. These practices are consistent with Cameron's recommendation that "empowering" researchers share knowledge with the researched and present the results in ways they and others who can benefit from them will find accessible.

Critical research places a different burden of ethical responsibility on the researcher than positivist research in which the researcher maintains an illusion of objectivity and distance. A particular danger associated with researching and writing about disadvantaged groups is that the researcher herself may contribute unwittingly to the oppression of the group by making statements which could be interpreted to support popular prejudices. Accordingly, I have become increasingly careful to avoid facile generalizations that could contribute to the popular stereotypes about men, women, and computers that have begun to surface in the mainstream media, such as the view that women send e-mail to socialize with friends while men send e-mail to exchange information (see e.g. Kantrowitz 1994). I even recently wrote an article debunking this stereotype (based on empirical linguistic evidence; Herring Forthcoming), something I would not have done several years ago. In short, critical research may call for considerably more follow-up than positivist research, and may actually lead one to take up an activist social agenda.

Critical CMC research is legitimate and necessary given the diverse and often conflicting needs of groups of users in cyberspace. However, the approach is incompatible with the suggestion that all CMC researchers seek consensus and approval from the researched as part of the research process. Requiring consensus would seriously limit and compromise the integrity of what critical research could be done, and thus is problematic on scholarly grounds. It is also theoretically naive, in that it

ignores the existence of multiple voices, multiple agendas, and struggle among the researched themselves. Finally, it risks reinforcing the hegemony of dominant groups, since theirs are the voices most likely to determine the form that any "consensus" will take. Thus although giving more of a voice to the researched may ease ethical concerns in the context of some research agendas, it is ethically problematic in others.

## SUMMARY AND CONCLUSION

I began this essay 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. 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.

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. How could any single set of guidelines hope to appropriately reflect the nature of the interaction in all of these different genres?

Fortunately, there is a simple solution, at least to ethical questions 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. In contrast, private arrangements must be explicitly set up and managed; on the Internet, these might include Listserv groups that require approval of the 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 group wishing to observe and report on electronic interaction: public interaction is repeatable for any reasonable and non-malicious use (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 source. The 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 a 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.<sup>6</sup> 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 (who clearly should convert their group to a private format, unless the publicness of their existing arrangement is therapeutic to them for other reasons -- in which case, they must decide whether they can live with the risk of exposure).

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. This kind of argument has been advanced to justify the need to respect the "perceived privacy" of individuals posting messages on open-access Usenet and Listserv groups. However, no one would attempt to make such an argument for television or radio broadcasts, because their public nature and the conventions of broadcasting are well understood by most adults in modern society. Such a widespread understanding of the public nature of CMC has not yet been achieved, yet the two phenomena are parallel in many ways. Thus if it is agreed that ethical guidelines are necessary, I propose that they be arrived at by clarifying and codifying understandings that are plausibly in the process of emerging; the default public nature of CMC is one such plausible understanding.

Underlying this proposal is an assumption that participants and listowners have a responsibility to themselves to protect their privacy. Researchers share in this responsibility, of course, given their more powerful position in the relationship, but researchers also have a responsibility to themselves and to their research. To expect them 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 protect researched populations from harm as a result of their research, in cyberspace as elsewhere. However, means for protecting the researched in cyberspace are already available and can be developed further, e.g. by exploiting the technical distinction between public (open-access) and private (restricted-access) groups. It is unnecessary to implement invariant guidelines which 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 diversity of research practices. This essay 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 years to come.

## Notes

1. 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.

2. This example is given in Chafe (1994:205).
- 3.
3. This doesn't 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.
4. I understand Fillmore's motives (in his 1974 paper) to be similar. He could very well have named the professor who posted (broadcast) the message on his office door. However this individual identity was not important to the linguistic point he was trying to make, and the example additionally made the professor look vaguely ridiculous. Hence the superficial disguise.
5. Approaching the listowner for permission is a possible alternative that is often proposed. However, this does not solve the problem of representativeness, nor the ethical problem of "giving the researched a say". If anything, to the extent that the researcher feels responsible or beholden to the listowner for granting his or her permission, the practice may create a bias in the research in favor of the listowner's agendas.
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 insure compliance, malicious failure to do so would need to carry professional and social stigma, and perhaps be legally actionable as well.

## References

- Adams, Peter; Alison Towns; and Nicola Gavey 1995 "Dominance and entitlement: The rhetoric men use to discuss their violence towards women." *Discourse and Society* 6(3):387-406.
- Cameron, Deborah 1992 "Respect, please!"; Investigating race, power, and language." In D. Cameron, E. Frazer, P. Harvey, M.B.H. Rampton; and K. Richardson, eds. 113-130.
- Cameron, Deborah; Elizabeth Frazer; Penelope Harvey; M.B.H. Rampton; and Kay Richardson, eds. 1992 *Researching Language: Issues of Power and Method*. London: Routledge.
- Cavazos, Edward A. 1994 "Intellectual property in cyberspace: Copyright law in a new world." In E. Cavazos and Gavino Morin (eds.), *Cyberspace and the Law: Your Rights and Duties in the On-line World*. Cambridge: MIT Press.
- Chafe, Wallace L. 1994 *Discourse, Consciousness, and Time*. Chicago: University of Chicago Press.

Collot, Milena and Nancy Belmore 1993 "Electronic language: A new variety of English." In J. Aarts, P. de Haan and N. Oostdijk (eds.), *English Language Corpora: Design, Analysis and Exploitation*. Amsterdam & Atlanta, Ga.: Rodopi. 41-56.

Dibbell, Julian 1993 "A rape in cyberspace." *Village Voice*, Dec. 21.36-42.

Dijk, Teun van, ed. 1985 *Handbook of Discourse Analysis*. London: Academic Press.

Gurak, Laura J. 1995 "The multi-faceted and novel nature of using cyber-texts as research data." In Teresa M. Harrison and Timothy D. Stephen (eds.), *Computer Networking and Scholarship in the 21st Century University*. Albany: SUNY Press.

Ferrara, Kathleen; Hans Brunner; and Greg Whittemore 1991 "Interactive written discourse as an emergent register". *Written Communication* 8(1):8-34.

Fillmore, Charles 1974 "Pragmatics and the decription of discourse." *Berkeley Studies in Syntax and Semantics* 1. Berkeley: UC Berkeley Department of Linguistics.

Foucault, Michel 1980 *Power/Knowledge: Selected Interviews and Other Writings, 1972-77*. C. Gordon, ed. Brighton: Harvester.

Harvey, Penelope 1992 "Bilingualism in the Peruvian Andes." In D. Cameron, E. Frazer, P. Harvey, M.B.H. Rampton; and K. Richardson, eds. 65-89.

Herring, Susan 1993 "Gender and democracy in computer-mediated communication." *Electronic Journal of Communication* 3(2). Special issue on Computer-Mediated Communication, ed. by T. Benson. Available from comserve@rpitsvm.bitnet. Reprinted in Rob Kling (ed.), *Computerization and Controversy*, 2nd edition. New York: Academic Press. (Forthcoming.)

1994 "Politeness in computer culture: Why women thank and men flame." In M. Bucholtz, A. Liang, and L. Sutton (eds.), *Communication In, Through, and Across Cultures: Proceedings of the Third Berkeley Women and Language Conference*. Berkeley: Berkeley Women and Language Group.

1996 "Posting in a different voice: Gender and ethics in computer-mediated communication." In C. Ess (ed.), *Philosophical Perspectives on Computer-Mediated Communication*. Albany: SUNY Press.

Forthcoming "Two variants of an electronic message schema." In S. Herring, ed. 77-101.

Forthcoming (ed.) *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*. Amsterdam: John Benjamins Publishing Company.

Herring, Susan, Deborah Johnson, and Tamra DiBenedetto 1995 "'This discussion is going too far!' Male resistance to female participation on the Internet." In M. Bucholtz and K. Hall (eds.), *Gender Articulated: Language and the Socially Constructed Self*. New York: Routledge.

Kantrowitz, Barbara 1994 "Men, women and computers." *Newsweek*, May 16:48-55.

Kiesler, Sara; D. Zubrow; A.M. Moses; and V. Geller 1985 "Affect in computer mediated communication." *Human Computer Interaction* 1:77-104.

Kim, Min-Sun and Narayan S. Raja 1990 "Verbal aggression and self-disclosure on computer bulletin boards." ERIC Clearinghouse on Languages and Linguistics, document ED334620. Washington, D.C.

Labov, William 1966 *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

Steiner, Linda 1989 "Feminist theorizing and communication ethics." *Communication* 12:157-173.

Van Gelder, Lindsey 1990 "The strange case of the electronic lover." In G. Gumpert & S. L. Fish (eds.), *Talking to Strangers: Mediated Therapeutic Communication*. Norwood, NJ: Ablex, 128-142.

Werry, Christopher Forthcoming "Linguistic and interactional features of Internet Relay Chat." In S. Herring, ed. 45-59.

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

Forthcoming "Oral and written linguistic aspects of computer conferencing: A corpus based study." In S. Herring, ed. 27-43.

Zimmerman, Don H. and Candace West 1975 "Sex roles, interruptions, and silences in conversations." In B. Thorne and N. Henley (eds.), *Language and Sex: Difference and Dominance*. Rowley, MA: Newbury House. 105-29.

To obtain a sample copy of The Information Society (TIS), choose one of the following:

Print the form below and mail to the Taylor & Francis Office nearest to you. In North America, send email to: sample-tis@taylorandfrancis.com Outside North America, fill out an electronic form that is available via Taylor & Francis' web site. Follow the path to "sample copy".

To purchase individual copies of The Information Society (TIS), visit the Taylor & Francis Journal Ordering Page. Single back issues are available for each journal, the price being obtained by dividing the institutional subscription rate by the frequency + 10%, which is currently about US\$42 (as of early 1999).

Contact Peggy Pagano, Journals Manager ppagano@taylorandfrancis.com OR 1-800-821-8312 (EXT.117) OR 215-269-0400, OR FAX: 215-269-0363.

Taylor & Francis retain a 2 year back stock of journals. Older volumes are held by our official stockists: Dawson (UK)Ltd, Back Issues Division, Cannon House, Folkestone, Kent CT19 5EE, UK to whom all orders and inquiries should be addressed. Tel: +44 (0) 1303 850101; Fax: +44 (0) 1303 850440.

To subscribe, clip the following form and mail to the address below:

-----  
THE INFORMATION SOCIETY  
Published quarterly, ISSN 0197-2243

- Please enter my institutional subscription at US\$140 (starting with vol 15)  
 Please enter my personal subscription at US\$69 (starting with vol 15)  
 Please send me a free sample copy

Payment options:

Check/Money Order Enclosed  
(please make checks payable to Taylor & Francis, US\$ only)

Please charge my:  VISA  MC  Amex  
Card # \_\_\_\_\_ Exp date: \_\_\_\_\_

Signature: \_\_\_\_\_

Telephone: \_\_\_\_\_

(required for credit card purchases)

or BILL TO: (please print)

SHIP TO (if different):

|                       |                       |
|-----------------------|-----------------------|
| Name _____            | Name _____            |
| Institution _____     | Institution _____     |
| Address _____         | Address _____         |
| City _____            | City _____            |
| State _____ Zip _____ | State _____ Zip _____ |

Mail this form to:

Taylor & Francis Inc.  
47 Runway Road  
Levittown, PA 19057  
Toll free- 1-800-821-8312 or  
Phone- 215-269-0400  
Fax- 215-269-0363

Outside the U.S. contact:

Taylor & Francis Ltd.  
Rankine Road  
Basingstoke, Hampshire  
RG24 0PR, United Kingdom  
tel: +44 (0) 256 840366  
fax: +44 (0) 256 479438