(Brennan, S. E., & Ohaeri, J. O. (1999). Why do electronic conversations seem less polite? The costs and benefits of hedging. *Proceedings, International Joint Conference on Work Activities, Coordination, and Collaboration (WACC '99)*, 227-235, San Francisco, CA.)

Why do Electronic Conversations Seem Less Polite? The Costs and Benefits of Hedging

Susan E. Brennan
Department of Psychology
State University of New York
Stony Brook, NY 11794-2500 USA
susan.brennan@sunysb.edu

Justina O. Ohaeri
Lucent Technologies
101 Crawfords Corner Rd.
Holmdel, NJ 07733
ohaeri@lucent.com

ABSTRACT

Electronic conversations often seem less polite than spoken conversations. The usual explanation for this is that people who are not physically copresent become depersonalized and less inhibited by social norms. While this explanation is intuitively appealing, we consider another possibility, based on the costs of producing "polite" utterances when speaking versus when typing. We examined a corpus of conversations generated by 26 three-person groups who interacted either face-to-face or electronically to do a collaborative memory task. We coded hedges (which mark an utterance as provisional) and questions (which display doubt or invite input from others), as people presented their own recollections, accepted, modified, or rejected those of others, and tried to reach consensus. Both of these devices are associated with politeness. For most people, hedging is more difficult when typing than when speaking because additional words are required, while marking an utterance as a question is equally easy in both media. The two groups made somewhat different use of these devices: Face-to-face groups hedged more than electronic groups, but both groups used questions just as often. We discuss how these and other differences emerge from the costs and affordances of communication media.

Keywords

Politeness, hedges, multimedia, communication, groups.

1 INTRODUCTION

When people work together in a group, they are concerned not only with the task at hand, but also with the manner in which they accomplish it. Interaction is guided by social needs, apparently universal ones, that have been described under the general rubric of politeness [5]. These needs include preserving one's own self-esteem or face, not threatening the face of others, not imposing, providing options to others, and showing solidarity. Social needs such as these place additional demands on individuals working in groups that do not arise when they work alone.

Politeness is an especially interesting issue when group members communicate electronically. With the popularity of email and teleconferencing, many have noticed that conversations conducted electronically often seem less polite and less inhibited than those conducted face-to-face [17]. Many people have experienced the phenomenon of "flaming" in electronic communication, and several explanations for this phenomenon have been considered [16, 17, 27]. For instance, with newer communication media, people may be unfamiliar with conventions for etiquette; without nonverbal cues, communication may be harder to coordinate, messages may be more ambiguous, and people may focus on messages rather than their interlocutors; and since interlocutors are not copresent, people feel depersonalized and anonymous, causing them to act uninhibited (that is, they cease to care about face-management needs).

But consider another explanation for differences in perceived politeness: the features of electronic communication media simply make it more difficult for people to serve face-management needs, although people may continue to *care* about these needs. To explore this hypothesis, we draw on the collaborative framework of language use developed by Herbert H. Clark and his colleagues (see [7]). In particular, having a conversation involves a number of activities, and each of these may incur effort on the part of participants; Clark and Brennan [8] describe these costs as *formulation*, *production*, *reception*, *understanding*, *start-up*, *delay*, *asynchrony*, *speaker change*, *display*, *fault*, and *repair* costs.

For instance, consider what happens in a face-to-face spoken conversation: while talking, a speaker can monitor an addressee's intonation and facial expression for feedback about whether the addressee is understanding or accepting what is being said. This makes it easy for the conversants to achieve a joint focus of attention and to come to the mutual belief that they are talking about the same thing. Speaking is relatively effortless for most people, so production costs are low. Speakers can rely on cues such as eye contact and intonation to figure out when it is their turn to speak. And adjacent turns can be assumed to be relevant to one another, which makes them easier to interpret.

But consider what happens when people converse using a teleconferencing program such as a "chat" program: Since they are not copresent, they must work harder to *ground* their utterances, that is, to make sure that they mutually understand. Typing is harder than speaking for most, and so producing an utterance takes more time. People cannot

WACC '99 2/99 San Francisco, CA USA

see one another's utterances being produced incrementally, but receive them only after the typist has hit carriage return, so any delay in the conversation is harder to account for (*is my partner still typing, is she waiting for my response, or has she left the room?*). Because of this ambiguity, two people may produce and send utterances in parallel, which can disrupt the relatedness between turns and the local coherence of the dialog.

The process of grounding in communication, then, suggests concrete predictions about how people balance such costs, and therefore, how their conversations will be shaped by different communication media. Thus, differences between face-to-face and electronic conversations may emerge more directly from the different affordances of these two media, rather than from users who are not copresent ceasing to care about the social aspects of doing the task.

We tested this idea using a corpus of task-oriented three-person conversations collected by Ohaeri [21]. The task was an activity that people often perform spontaneously and interactively, in small groups–collaborative remembering. By collaborative remembering, we mean reminiscing together to fill in the details of an event that members of a group have witnessed together and coming up with a joint account of the event. Consider what they need to do: Each member of the group needs to recall what they believe matters for current purposes, contribute this information to the conversation in a timely and relevant manner, and do so in a way that lets the rest of the group know how confident they are that the information is accurate. Also, they must evaluate the contributions of others, and if these differ from their own recollections, they must negotiate a version they can all agree on. And they must do all this without losing face and without offending or threatening one another, or else consensus may become impossible.

Speakers have many metalinguistic devices at their disposal for displaying their commitment to what they are saying [4, 28]. For instance, when they are not entirely sure of the answer to a question, they may pause before producing the answer, fill that pause with an interjection such as *uh*, or produce the answer with rising intonation. On the other hand, when they believe they know the answer but just can't produce it, they may pause longer or say *uh* before answering *I don't know* [4, 28]. Furthermore, hearers are able to interpret this information correctly [4]. And when there are numerous perspectives that can be taken on an object, a speaker is more likely to mark a reference to it by hedging, as if to show that the referring expression is only provisional [2]. With evidence that an addressee finds the referring expression acceptable, the hedges drop out in subsequent references [2].

In their treatise on politeness, Brown and Levinson [5] present hedges and other devices in terms of people's needs to save face or avoid threatening the face of others. These needs are especially real when people find themselves disagreeing with one another in the process of reaching a consensus. However, we will point out (and Brown and Levinson would probably agree) that speakers do not apply these devices uniformly or indiscriminately to their speech in order to be polite, but that these devices bear precise information about the speaker's commitment to particular propositions and about their willingness to have this information modified by a partner. The ability to mark utterances as more or less provisional is a useful tool for reaching consensus.

With hedges and in other ways, speaker display their alignment to what they and their partners say [6, 13]. When one member of a group makes a proposal, she may frame it confidently, perhaps with an enthusiastic positive evaluation; or she may frame it more cautiously, by marking it with a hedge, rising intonation, hesitation, or interjection such as um. Other members may react to her proposal in a variety of ways: They can accept it explicitly with an acknowledgment or by repeating all or part of it. They can accept it implicitly by going on with the next relevant installment. They can disagree with it more or less forcefully (*No, I don't think so*; or *Do you really think so?*) or modify it explicitly (*No, it was X*; *or I think it was X, not Y*; or *But wasn't it X?* or implicitly (with *um* or *well*). They can even ignore a proposal entirely. Cahour and Pemberton [6] describe this alignment process in terms of whether a speaker appears to associate or distance herself from the utterance she is presenting, as well as by whether her addressee, by his response, appears to associate or distance himself from the utterance.

By displaying their alignment toward their own and their partners' messages, speakers are able to achieve true politeness, which results from giving a partner options. True politeness differs from conventional politeness, which involves merely adding words like *please* and *thank you* or else replacing a direct command with an indirect construction such as *May I ask you to…* (see [3] for examples of the effects of conventional politeness in human-computer interaction). In a task that requires reaching consensus, distancing oneself from one's own utterance may make it easier for one's partner to offer a counterproposal and still preserve everyone's face.

Coordination between speakers has been studied extensively on the level of how speakers ground meaning during referential communication (e.g., [1, 2, 10, 15, 18, 26]). One advantage of studying communication using referential communication tasks is that these tasks include a built-in measure of the success of communication. Such studies capture how people are able to achieve a mutually shared focus of attention on the same object, a necessary and basic activity. But referential communication tasks may not provide the best arena for tapping some of the other important achievements that people manage routinely in conversation, such as convincing one another that certain information is reliable or reaching agreement about something that may evoke diverse opinions or emotional responses. Studies of interaction that use open-ended brainstorming or design tasks provide a bit more room for such activities (e.g., [6, 11, 12]). But the problem with using open-ended tasks is that there is no objective measure of success, particularly when consensus is not required.

Our task, collaborative remembering, combines the advantages of referential communication and brainstorming: there is an objective standard to compare with the group's recall product, but it is a more complex social task than simply picking out objects. The component activities of

the task serve a more or less coherent whole, and there is plenty of room for interpretation and disagreement. And most important, it is something that people do frequently in everyday interaction. Our goal, then, is to examine the alignment process in collaborative remembering by looking at the use of hedges and other devices, to see how this may differ between face-to-face (spoken) and electronic (typed) conversations.

2 THE STUDY

The corpus we used was generated for a larger study of how group processes influence individuals' memories for events (see [21]). In the larger study, groups of three people watched a movie clip, recalled it alone (Session 1), recalled it together with the rest of the group (Session 2) with the requirement that they reach consensus on the story, and then recalled it alone once again (Session 3). For the present study, we focused only on Session 2, during which groups met either face-to-face (13 groups) or electronically using a chat program (13 groups).

2.1 Participants

Seventy-eight undergraduate psychology students at the State University of New York at Stony Brook volunteered to participate in exchange for research credit. All were native speakers of English and were able to type; all reported never having seen the movie from which the story was drawn.

2.2 Materials

The groups watched an 8-minute clip from a film by John Sayles, The Secret of Roan Inish. In the clip, an old man tells a story to his granddaughter. The film cuts back and forth between the telling of the story and the story itself. Recalling an event from a film was chosen because it is an activity familiar to most people. The particular clip was chosen because it is engaging, short enough not to be overwhelming, and contains a complete story with a relatively simple structure.

Face-to-face conversations were recorded using an audio tape-recorder and stereo microphone. Electronic conversations were conducted over networked Macintosh computers running Aspects™, a chat software program by Logic Technologies. A chat window and a text document appeared side by side, extending the full length of the monitors; these were visible to all three group members. The chat window was used for idea presentation and discussion and indicated which statement came from which group member (labeled as A, B, or C). The text document was used to record the group's final product. At the bottom of the chat window was an edit window where participants typed their contributions to the conversation. This was the only space that was not shared; it was visible only to the person typing in it. Messages were sent to the chat window by hitting the 'Return' button. All editing of messages was done in the edit window; messages already in the chat window could not be edited. The text document for the group's product was fully editable (but by only one of the group members).

2.3 Procedure

Groups of three people watched the movie clip together, after which they did a 5-minute distracter task individually (listing the names of as many countries as they could think of). Then they were taken to separate rooms for the first recall session, where each recalled the movie alone while typing into a text editor. For the second session, the one we focus on here, groups were assigned on an alternating basis to either the face-to-face or electronic condition.

Face-to-face groups were seated together near the same computer so that all could see the monitor (they could see each other as well). They were instructed to discuss the movie clip and come up with a consensual account of it. One member was randomly assigned to type up the group's account; this person also took part in the discussion.

Electronic groups were brought together around the same computer and trained to use the chat software by typing into the edit window and then pressing return to send the message to the chat window. After this training, they returned to their separate rooms. As in the face-to-face condition, they were instructed to discuss the movie and reach a consensus and one member was randomly assigned the task of compiling the group product on the shared text document. Although all three members could see and comment on what was written on this document, only the assigned person was able to write to it. Transcripts of the session (both chat and text windows) were saved into a text file.

Two hours were allotted for the entire experiment. Within this period, time for group sessions was not limited, in order to accommodate the additional time we expected would be required for typing in the electronic condition. Although we did not formally time the sessions, we observed that those in the face-to-face group took about 90 minutes to complete the three recall sessions, and those in the electronic group took about 10-15 minutes longer.

3 CODING AND DATA ANALYSIS

3.1 Transcripts

The audiotapes of the face-to-face sessions were transcribed and checked.[1] Word counts and turn counts were computed for each group session. Excluded from these counts were nonverbal sounds like laughter, as well as any speech that was unintelligible. Turn beginnings were counted whenever there was a speaker-change in the conversation. The electronic group sessions yielded chat transcripts organized into turns labeled with each group member's identifying letter.

Transcripts were broken down into propositions. A proposition was defined as an independent idea unit that contributes new information. In addition to phrases that told the story, descriptive adjectives, temporal and spatial locatives, and quantities were coded as separate propositions. Proper names or other references to individuals, as well as adjectives describing a particular person or object, were counted at first mention only (that is, expressions that re-referred to the same entity without adding information were not counted). Direct and indirect quotations were broken down into constituent propositions.

[1]Transcriptions included exact wording but did not capture overlapping speech, as this level of detail was not relevant to our study.

For each transcript, the total number of propositions was tallied, as well as the numbers of propositions representing correct and incorrect details of the story.

3.2 Markers of provisionality

3.2.1 Hedges

To capture how speakers align themselves with their own and each others' utterances, we coded two types of markers. The first type is hedges. When hedges modify the speaker's own utterance, they display that the speaker is taking a provisional stance toward the utterance; they may grant the addressee license to reject or modify the utterance. When they follow another's utterance, they may either cast doubt or add cautious support (probably depending on how committed the other speaker appears to be). Note that the interactive options provided by hedging appear to serve both task-related and face-management needs. We coded a number of different sorts of hedges, such as when recalled information was marked by expressions such as *kind of*, *whatever*, *something*, or *like* (coded instances are underlined):

> Yeah, they were sitting around the fireplace in the night... sort of like a bedtime story kind of thing. *[face-to-face group 4]*

Sometimes speakers modified their commitment to a particular term by adding the morpheme *-y*, and we counted such instances as hedges, as here:

> We all agree it was a wreathy thingy on his neck??? *[electronic group 8]*

We also coded as hedges expressions of doubt that modified recalled information such as *I think, I guess*, and *I don't know*:

> We all agree it was in Ireland, I think? *[face-to-face group 3]*

We did *not* count claims of forgetting in general or claims that modified information that was not recalled, such as:

> Did they give the name of the island? 'cos I don't remember. *[face-to-face group 4]*

3.2.2 Questions

Dialogs are structured into adjacency pairs [20, 25] where a first part, such as a question, projects a second part, such as an answer. We were interested in how people pursue responses from their partners [22], and so the second group of markers we coded were questions. The questions we coded included syntactic questions (beginning with a wh-question word or auxiliary verb) and utterances marked as questions by rising intonation (face-to-face) or punctuated by question marks (in the electronic medium). They also included tag questions, such as these:

> He was cold and almost dead... the women tied him up between two cows... he started to sweat, right? *[face-to-face group 4]*

> M: He began telling the story
>
> S: about when his father was young, when Ireland was being ruled by the British
>
> A: It still is, isn't it? *[face-to-face group 10]*

We make two comments before turning to our results: Hedges are not "polite" in and of themselves. Even though we call the hedges and questions that we have coded "markers" of provisionality, we do not mean to imply that, by their very presence, they enable people to automatically achieve politeness in a signaling-type manner (see, for instance, Levinson's objection [20] to Labov & Fanshel's theory of "mitigators" [19]). Instead, we follow the view of politeness that Brown and Levinson [5] developed within Grice's cooperative framework [14, 15].[2] On this view, people experience an interaction as more or less polite due to implicatures they make based on the pragmatic features of the situation. True politeness gives a partner options. What hedges and tag questions do signal is that a speaker's stance toward a proposition is one that invites (but does not demand) a partner's input. To the extent that this serves the face-needs of a group, the interaction is experienced as polite.

4 RESULTS AND DISCUSSION

4.1 Performance

Most people find it easier to talk than to type; consistent with this, face-to-face groups produced nearly twice as many words as electronic groups, $t(24) = 3.79$, $p < .001$. Despite this difference, electronic groups recalled just as many propositions from the story as face-to-face groups (see Table 1). And the quality of these memories was no different for the two kinds of groups: Face-to-face groups made errors 5% of the time, and electronic groups, 6% (not significantly different). So groups performed equally well on the task of recalling the story, regardless of the communication medium.

4.2 Hedges

We found that people marked their contributions with hedges much more often when they were talking (2.5 hedges per 100 words) than when they were typing (1.2 hedges per 100 words), $t(24) = 3.52$, $p < .005$. Hedging,

[2] On the other hand, we do not adopt Brown and Levinson's pessimism about whether the strategies for satisfying face needs can be quantified. Our goal is to understand the interactional devices people have at their disposal for reaching consensus. We believe that much can be learned by comparing the distributions of linguistic events (or "markers") over two kinds of situations that vary systematically (as with our face-to-face and electronic conditions).

which appears to signal a speaker's level of commitment to an utterance, happens more often in spoken than typed utterances. Why should this be? Showing one's alignment to an utterance by hedging it has a significant cost, particularly when producing the utterance is effortful (as typing is for many people). This increased cost lowered hedging rates overall in electronic groups. But within the 13 electronic groups, hedging should be more common for those whose members type easily or who do not mind typing, than for those for whom typing is a chore. We did not measure group members' typing speed directly; however, we can estimate a group's overall aversion to typing by looking at their total word count. Electronic groups produced word counts that ranged widely, from 505 to 1348 words. And there was a significant positive correlation between word counts and hedge rates, $r(13) = .55$, $p < .05$. If people were behaving less politely because of depersonalization caused by the electronic medium, we would not expect to find this correlation. Furthermore, this correlation was present only for electronic groups–in face-to-face groups, word counts and hedge rates turned out to be independent, $r(13) = .17$, *n.s*. While many electronic exchanges certainly appear to be less polite than many face-to-face exchanges, our data support the hypothesis that this is due to the added cost of deploying devices such as hedging, rather than any general depersonalization caused by the medium.

Table 1. Results for Face-to-Face v. Electronic Groups

	Medium		
	Face-to-Face	Electronic	*Different?*
# of Words	1646	886	$p < .001$
# of Turns	199	96	$p < .001$
Turn Length *(In Words)*	8.3	9.6	$p < .10$
Total # Recalled Propositions	79.7	74.9	*no difference*
Errors	3.7 (5%)	5.1 (6%)	*no difference*
Hedge Rate *(Per 100 Words)*	2.5	1.2	$p < .001$
Question Rate *(Per 100 Words)*	2.5	2.5	*no difference*
Acknowledgments *(Per 100 Words)*	1.5	.9	$p < .001$

4.3 Questions

Consider the cost of pursuing a response from an addressee or marking one's commitment to an utterance by using wh-questions, tag questions, or try markers. When face-to-face, speakers can rely on syntactic or intonational means to frame an utterance as a question. When typing, they can also rely on syntactic means or on the relatively painless device of punctuating an utterance with a question mark. The devices for pursuing a response, then, do not appear to be significantly more costly electronically than face-to-face. Indeed, we found that electronic groups were just as likely to use questions as face-to-face groups, 2.5 times per 100 words, $t(24) = .15$, *n.s.* Taken together with the hedging results, this provides support for the idea that the costs of communicating in various media shape the discourses people produce in those media.

4.4 Discourse Structure

How did face-to-face groups manage to produce the same number of propositions in twice as many words? For one thing, these groups tended to ground somewhat smaller installments than did electronic groups, and they often did so with explicit acknowledgments (*yeah*, *yes*, *right*, *mhm*, *ok*, *uh huh*, *true*, *sounds good*, *I agree*) or by repeating parts of a previous utterance. Consider the many examples of explicit and implicit acknowledgments in this example from a group conversing face-to-face:

> A: yeah, and that the kid was like talking, something
>
> L: yeah he was talking that's why he got in trouble he was whispering in some kid's ear about something?
>
> A: yeah
>
> L: um... then he gets punished or whatever?
>
> D: what was that, a wreath or
>
> L: yeah it was some kind of browny
>
> A: yeah it was some kind of straw thing or something
>
> L: mhm
>
> D: around his neck
>
> L: so that everybody knew what he did or something?
>
> A: straw wreath
>
> D: yeah
>
> A: and then
>
> L: and then they went outside, right
>
> A & D: yeah
>
> L: and all the kids were like taunting him and stuff?
>
> D: yeah... and then he saw the teacher and he just went after him
>
> L: he first took the thing off?
>
> D: mm... did he punch him in the face?
>
> L: yeah he was punching him and he was cursing at him as well
>
> D: yeah
>
> *(laughter)* *[face-to-face group 13]*

Although explicit acknowledgments were also used by electronic groups, they were used more sparingly than by face-to-face groups ($t(24) = 5.30$, $p < .001$). Often, rather than acknowledging explicitly, one member accepted another's contribution implicitly, by simply going on to present the next relevant part of the story. For example:

> C: One day the boy gets caught talking to a classmate. He has to wear a wreath around his neck as punishment.

B: All of his classmates lauh *[sic]* at him and make fun of him. The boy gets so mad that he throws off the wreath and beats up the school master

A: he takes it until he can't take it anymore

[electronic group 12]

This pattern of grounding larger constituents while typing and smaller constituents while speaking is consistent with Clark and Brennan's [8] predictions about grounding in different media. Because, in our study, people constructed typed utterances in private areas before sending them to the public chat window, and messages could be produced in parallel, electronic groups could not assume they would all be attending to the same details of the story at any given moment, nor could they assume a turn would be relevant to the one before it in the chat window. That they often presented more finished or elaborated ideas, then, is not surprising. On the other hand, since people were copresent in the face-to-face conversations, they could more reliably determine if and when they shared a focus of attention; also, speaking was relatively easy, as was changing speakers; and the construction of one speaker's utterance was witnessed by all group members in real time. Therefore, they were in a better position to collaborate on smaller constituents. Consistent with this, our face-to-face groups produced marginally shorter turns than our electronic groups, 8.3 to 9.6 words, $t(24) = 1.74$, $p < .10$.

Although we did not attempt to quantify the amount of repetition that occurred in the conversations, inspection of the transcripts shows more repetition in face-to-face than electronic conversations (see the two examples just presented). This pattern is also consistent with the different costs of grounding utterances in the two media. That is, electronic group members can easily review the conversational record, because it is preserved in the chat window. They can also assume that this record is continually available to other members of the group, and so there is no need to repeat material. But in face-to-face conversations, speech is ephemeral. By repeating all or part of a previous utterance, a speaker ensures that the necessary information will be available in working memory, and that partners will be able to tell just how an utterance is relevant to previous discourse.

Another aspect of discourse structure that showed interesting differences was the organization of topics, and this can be explained by the features of the medium as well. In spoken conversations, turns adjacent in time are ordinarily relevant to one another [8, 20, 25]. But in our electronic condition, since typing was relatively slow, producing a timely response was more difficult, and turns often were relevant to material that appeared several turns back. Members could not witness each other's messages being composed, and so they often composed messages in parallel. As a result, although messages arrived adjacently in the chat window, they were often on different topics. Topic structure was further disrupted because people often responded to several previous topics in a single turn. In this way, many electronic conversations developed multiple "threads." The next example illustrates this point.

B: Then go on to tell about how the boy (shall we name him) went to the school were *[sic]* they couldn

B: sorry hit return . couldn

A: ok-I got you

C: The ol woman is the grandmother

A: OH yeah- what is she doing with that yarn?

B: did it again. where the boy couldn't speak Irish and when he was caught we punished by putting that thing around his neck. yeah she's grandma. I think she's making a doll

C: She is knitting a sock or a sweater

B: she is knitting something and leave it at that

A: How about doing a craft

B: good idea

C: No that's not why they punished him. It was because he was talking in class

B: no he was speaking Irish and it wasn't allowed

B: what do you say number 1

C: What do you mean. I think I missed that part of the clip

A: I think he wasn't supposed to speak Irish within an "easrshot *[sic]* of the teacher

B: yeah *[electronic group 9]*

4.5 Managing Disagreement

The transcripts contained an interesting range of ways in which people managed disagreements. Occasionally someone introduced a dissenting opinion straightforwardly, with terms such as *no*, *well*, and *actually*, as in this example:

C: ok

A: no, the sea got angry when he was rescued

C: ok

B: oh *[electronic group 11]*

This example seems less polite than the next one (also from an electronic group), in which all three members hedge abundantly about whether the story's main character's hands were bloody or not. Participant A settles the matter by declaring that the character's hands were bloody but not covered with blood; her use of *no* actually shows solidarity, since she is contradicting B's own hedge that he could be wrong.

B: something about his hands being bloody too

A: When were his hands bloody?

B: after beating the teacher

C: I dont remember his hands being bloody - maybe

B: i could be wrong

A: No I thinkj *[sic]* you're right. THey weren't covered or anything

C: they wernt covered with what

A: blood *[electronic group 2]*

This example shows that groups who are facile with the electronic medium (such as electronic group 2, who, in their chat session, produced the second-highest number of words of any electronic group) do manage to use devices such as hedges. To the extent that hedging provides other group members with options so that they can disagree without losing face, disagreements can be resolved with politeness.

5 CONCLUSIONS

We have examined just a few of the ways in which people can align themselves with their own and their partner's utterances, in the service of managing face needs. When a speaker hedges a statement, she opens the door for her partners to modify it, so that no one is likely to lose face. The act of hedging often requires modifying an utterance with additional morphemes, words or phrases. This is relatively easy for speakers, but requires extra effort from most typists. Consistent with this cost-based account, our face-to-face groups hedged more often than our electronic groups. However, those electronic groups that were not averse to typing produced hedges at a greater rate than those who typed less, approaching the rates of hedges by speakers in the face-to-face groups. We would not expect this if the differences between the two kinds of groups were due simply to depersonalization in the electronic groups.

Although people in electronic groups used hedges and acknowledgments less often and grounded somewhat larger constituents, they still managed to display their alignment to their own and their partners' utterances in other ways. The costs of marking an utterance as a question appear to be roughly equivalent in the two media, and electronic groups used questions (syntactic questions, try markers, and tag questions] just as often as face-to-face groups. We take this as evidence that people communicating via text still cared about face-management needs, and when hedging rates were low, this was because of the effort of typing. People allocated more of their effort to the message because typing was harder than talking. In this task we found no evidence that people using a chat program ceased to care about social tasks such as face-management or were deindividuated.

Our results are consistent with other studies that have found electronic communication to be less efficient for reaching consensus than face-to-face conversations (e.g., [27]). They are also consistent with the idea that people communicating electronically focus more on the message than on the social context. We expect that this happens because the primary task (in our experiment and in many other situations) is often more directly related to a non-social than a social objective. We expect that people would distribute their efforts differently if the primary objective were to socialize rather than to recall a story.[3]

We believe that our results are due primarily to the different costs of remote, typed conversation versus face-to-face, spoken conversation. If this is the case, then our findings about textual exchanges should generalize to email exchanges as well, particularly to extremely interactive email exchanges that take place over a short time (with little or no editing). We had the electronic groups use a chat program rather than email in order to minimize turn-taking costs and make the electronic medium as similar as possible to spoken conversation.

Our findings hold several implications for computer-supported cooperative work and computer-mediated communication. First of all, if people are aware of how the features of a medium interact with their tasks, they may be able to compensate for any undesired effects, such as text messages that seem brusque or rude. An experienced user may choose to devote more time to producing and editing messages in order to explicitly display her alignment toward an issue or a message. As the recipient of an ambiguous message, she may learn not to assume that the message was intended to be offensive. Although we do not expect that perceived politeness differences in electronic communication would disappear entirely with experience, experience does appear to help (consider, for instance, the evoluation of new expressive conventions, such as horizontal faces constructed out of punctuation, :-), to show that a remark is intended humorously). A second implication is that if several media are available for an intended conversation, the choice of which one to use should depend on factors such as whether the point of the conversation is to reach consensus. Third, the designers of CSCW systems should be sensitized to the effect that a seemingly trivial feature may have on the form and experience of a conversation. An experience such as "politeness" is not simply a style of interaction that emerges automatically in a medium. To the extent that people can be helped to manage the additional costs of displaying their alignment toward their utterances in mediated communication, both task and face needs will be served.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. IRI9402167. Justina O. Ohaeri was supported by a W. Burghardt Turner Fellowship. We are grateful to Ricardo Carrion, Tara Defantis, Christina Paul, and Amanda Pisino for their

[3]Note that we do not claim that *all* differences between spoken and text conversations can be explained in terms of effort, or that text-based conversations are never depersonalizing. In their study, Siegel et al. [27] found more swearing, name-calling, and insults in computer-mediated conversations than face-to-face. It could be that problems escalate more rapidly in electronic communication because people have not increased the amount of effort they spend on face-management goals, and *over the course of the conversation*, they become depersonalized or uninhibited.

assistance in running the experiment, transcribing, and coding, and to Michael Schober for helpful comments on a draft of the paper. Please address correspondence to Susan E. Brennan, Department of Psychology, State University of New York, Stony Brook, NY, 11794-2500

REFERENCES

1. Bortfeld, H., & Brennan, S. E. (1997). Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes, 23*, 119-147.

2. Brennan, S. E. & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 6*, 1482-1493.

3. Brennan, S. E. and Ohaeri, J. O. (1994). Effects of message style on users' attributions toward agents. *CHI '94, Human Factors in Computing Systems,Conference Companion,* pp. 281-282. Boston, MA.

4. Brennan, S. E. and Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language, 34*, 383-398.

5. Brown, P. & Levinson, S. C. (1978). *Politeness: Some universals in language usage.* Cambridge, UK: Cambridge University Press.

6. Cahour, B. & Pemberton, L. (1998). When people design: Conversational positioning and roles in collaborative design dialogues. *Proceedings, COOP '98.* Cannes, France.

7. Clark, H. H. (1992). *Arenas of language use*. Chicago: The University of Chicago Press.

8. Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L.B. Resnick, J. Levine, & S.D. Teasley (Eds.), Perspectives on socially shared cognition. Washington, DC:APA. Reprinted in R. M. Baecker (Ed.), *Groupware and computer-supported cooperative work: Assisting human-human collaboration.* San Mateo, CA: Morgan Kaufman Publishers, Inc.

9. Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13,* 259-294

10. Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1-39

11. Dennis, A. R., & Valacich, J. S. (1993). Computer brainstorms: More heads are better than one. *Journal of Applied Psychology, 78*, 531-537.

12. Gallupe, R. B., Bastianutti, L. M., & Cooper, W. H. (1991). Unblocking brainstorms. *Journal of Applied Psychology, 76*, 137-142.

13. Goffman, E. (1981). *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

14. Grice, H. P. (1957). *Meaning*. Philosophical Review, 66, 377-388.

15. Grice, H. P. (1975). Logic and conversation (from the William James lectures, Harvard University, 1967). In P. Cole, & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts*. (pp. 41-58). New York: Academic Press.

15. Isaacs, E. A., & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General, 116*, 26-37.

16. Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist, 39*, 1123-1134.

17. Kiesler, S., Zubrow, D., Moses, A., & Geller, V. (1985). Affect in computer-mediated communication: An experiment in synchronous terminal-to-terminal discussion. *Human-Computer Interaction, 1*, 77-104.

18. Krauss, R. M., & Glucksberg, S. (1969). The development of communication: Competence as a function of age. *Child Development, 40*, 255-256.

19. Labov, W., & Fanshel, D. (1977). *Therapeutic discourse: Psychotherapy as conversation*. New York: Academic Press.

20. Levinson, S. C. (1983). *Pragmatics*. Cambridge, UK: Cambridge University Press.

21. Ohaeri, J. O. (1998). Group processes and the collaborative remembering of stories. *Doctoral Dissertation*. State University of New York at Stony Brook.

22. Pomerantz, A. (1975). Pursuing a response. In J. M. Atkinson and J. Heritage (Eds.), *Structures of social action*, pp. 152-163. Cambridge: Cambridge University Press.

23. Sacks, H., & Schegloff, E. (1979). Two preferences in the organization of references to persons in conversation and their interaction. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology*, pp. 15-21. New York: Irvington Publishers.

24. Sacks, H., Schegloff, E., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language, 50*, 696-735.

25. Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica, 8*, 289-327.

26. Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology, 21*, 211-232.

27. Siegel, J., Dubrovsky, V., Kiesler, S., & McGuire, T. (1986). Group processes in computer-mediated communication. *Organizational Behavior and Human Decision Processes, 37*, 157-187.

28. Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language, 32*, 25-38.