The Collaborative Model and Psychological Evidence: Some Principles for Designing Cooperative Systems

Dale BarrBoaz Keysardaleb@ccp.uchicago.eduboaz@ccp.uchicago.edu

The University of Chicago Department of Psychology, 5848 S. University Ave., Chicago IL 60637 daleb@ccp.uchicago.edu

ABSTRACT

The collaborative model developed by Clark and colleagues views language use as a process by which people collaborate to perform joint actions, not merely exchange information. Because language use is one fundamental way in which people engage in cooperative interaction, many aspects of the collaborative model might be fruitfully applied to the design of computer systems which support cooperative interaction between computers and their users, or among humans. We highlight those aspects of the collaborative model which appear most applicable to the development of cooperative systems: reference diaries, optimal design, grounding, and signaling. In each case, the model has implications for system design. However, we review evidence which suggests that language users do not typically behave in all manners predicted by the collaborative model. We attempt to synthesize the predictions of the collaborative model and empirical evidence to derive a set of design principles to optimize computer/user interactions.

Keywords

Common ground, psychological evidence, grounding, signaling, perspective adjustment.

1. INTRODUCTION

People seek to coordinate their actions in order to accomplish individual and collective goals. Language use is probably the primary means by which such coordination is achieved. With the recent emergence and spread of the Internet and multi-media communications technologies, computers have become a common site in which people use language to coordinate action. In order for computer-mediated coordination to be effective, computers should be designed with the nature of their users in mind: they should reflect what people need, desire, and expect when they attempt to coordinate action through language. The work of Herb Clark and colleagues, which we will refer to as the "collaborative model" of language use, provides one possible "theory of the user." Many helpful guidelines for the design of computer systems can be derived through the application of the collaborative model. In what follows, we seek to identify those application areas which seem particularly promising for the designer. For each of these areas, we will evaluate the claims of the model by reviewing the available psychological evidence. Finally, we will attempt to synthesize the collaborative model with the empirical findings in order to derive a set of design principles.

2. COMMON GROUND AS THE CONTEXT OF LANGUAGE USE

The collaborative model regards language as a medium by which individuals coordinate joint projects. A distinguishing feature of the collaborative model is its emphasis on common ground. Common ground is defined as the sum of the "mutual knowledge, mutual beliefs, and mutual suppositions" shared by individuals (Clark, 1992, p. 3). In order for a proposition p to be included in the common ground of two people, it requires more than just that each person knows that p, but also that they mutually know that p. According to Clark and Marshall (1981) people seek evidence that information is part of common ground by using "co-presence heuristics." Two people can assume that a piece of information is part of common ground if: 1) it is physically or perceptually present to them (physical co-presence); 2) it has been mentioned previously in conversation between them (linguistic co-presence); or 3) it is common knowledge in one of the communities of which they are both members (community membership).

2.1. Reference Diaries

According to the collaborative model, people maintain specialized memory structures known as "reference diaries" to keep track of their common ground with others (Clark & Marshall, 1981). Not only do people remember facts, but they also remember with whom these facts are mutually shared. For example, when Ann and Bob bump into each other at the cinema on the way out from the film Titanic, not only will Ann encode the fact that she saw the movie, but also that Bob saw it. A week later, Bob calls Ann and asks, "What did you think of the movie?" If Ann consults her reference diary for Bob, she should understand that Bob is referring to Titanic and not to Cinema Paradiso, which she had seen without him.

2.2. Optimal Design

The claim that people manage information in the form of reference diaries makes specific empirical predictions about memory encoding and retrieval during language use. Reference diaries provide a knowledge base from which people formulate and interpret messages with respect to particular addressees. According to the collaborative model, when people design an utterance, they should use only the information included in the common ground with their addressees. This assumption is embodied in the principle of optimal design, which states that a speaker "designs his utterance in such a way that he has good reason to believe that the addressees can readily and uniquely compute what he meant on the basis of the utterance along with the rest of their common ground." (Clark, Schreuder, & Buttrick, 1983, p. 246). Optimal design is also assumed to play a role in message comprehension, because addressees expect that speakers obey this principle. Therefore, addressees would understand a message optimally if they access exclusively the common ground information stored in their reference diary for the speaker: "...the comprehension process must keep track of common ground, and its performance will be optimal if it limits its access to that common ground." (Clark & Carlson, 1981, p. 328).

If it is true that people store information in reference diaries, that they follow the principle of optimal design in formulating utterances and comprehend utterances in the context of common ground, then we might derive the following guidelines for the design of a computer system. Computer systems which are optimized to support cooperative interaction should store information in a user-specific manner, and constrain information search to only that information which is shared with a particular user or set of users. Otherwise, users will encounter problems when presented with information which violates common ground. For example, when an average user executes a command and receives a message which specifies technical information for programmers to use in debugging, this information will only serve to confuse the user about the outcome of his or her action, because he or she is not a member of the community of experts for whom this message is meaningful. If the computer stored information in the form of reference diaries, it could custom-tailor its output to conform to the general background and expertise of particular users. This would permit the system to formulate and interpret user-messages in a manner which is "optimal" according to the collaborative model.

However, the suggestion that common ground and mutual knowledge play such an extensive role in language use was challenged early on (e.g., Johnson-Laird, 1982; Sperber & Wilson, 1982). More recently, empirical data concerning how people actually use common ground in conversation have not supported the claims of the collaborative model in two important ways: 1) several studies which appear to demonstrate the use of common ground yield to simpler alternative explanations, due to a confound in their design; and 2) studies which correct this design problem indicate that people often do not meet the collaborative model's standard of optimality.

Keysar (1997) analyzed the design of several key papers which sought to test the hypothesis that people use common ground in comprehension (e.g., Clark, Schreuder & Buttrick, 1983; Greene, Gerrig, McKoon & Ratcliff,

1995). The analysis revealed that the design of these studies systematically confounded common ground with personal salience. Although these studies demonstrate that people used an available piece of information to comprehend utterances, they did not show that they used this information because it was mutually shared. In fact, even when people have informed and accurate reference diaries—when they are certain about who knows what—they don't always use them in communication. Keysar (1994) showed that when readers assess how an addressee would understand an utterance they often rely on information which they know is inaccessible to that addressee. When readers know the intention of a speaker, they tend to believe that addressees perceive that intention even when they know that the addressee is missing crucial information.

Furthermore, there is little evidence that the principle of optimal design plays a role in language processing, in that speakers and comprehenders do not appear to use it in formulating or interpreting utterances. Rather, studies have shown that common ground is used mainly as a correction mechanism. Horton and Keysar (1996) showed that speakers initially plan their utterances independently of common ground, but use common ground to monitor their plans and revise them if they violate the perspective of the addressee.

Just as speakers appear to use common ground to correct erroneous plans, addressees appear to use common ground to correct errors of understanding, but not to restrict their initial interpretation. In several studies, we have examined how addressees understood utterances by tracking their eye-movements while they interpreted a speaker's referential descriptions. We found that addressees consider information which is salient to them, even if they know that it is inaccessible to the speaker (Keysar, Barr, Balin & Brauner, 1996; Keysar, Barr, Balin & Paek, in press; see also Hanna, Trueswell, Tanenhaus & Novick, 1997). We take these results to indicate that language use has a large egocentric component (Keysar, Barr & Horton, in press).

2.2. Optimality and Perspective Adjustment

These results do not accord with many people's intuitions about language use. Why shouldnÕt speakers and addressees use their knowledge of what is commonly known, especially when egocentrism could lead to systematic miscommunication? According to the collaborative model's view of how people should behave, the egocentric behavior revealed in empirical studies looks suboptimal. In contrast, the perspective adjustment model of language use (Horton & Keysar, 1996; Keysar, Barr, Balin and Brauner, 1996; Keysar, Barr, Balin and Paek, in press) views the egocentric behavior of language users as "optimal", not because it ensures fail-safe communication, but because it simplifies the task of producing and comprehending utterances. According to the perspective adjustment model, people use language egocentrically as a heuristic to reduce its ambiguity. Because common ground is a kind of meta-knowledge which is cognitively effortful and time-consuming to access, people will use it only to monitor production or comprehension processes for errors, but not when they formulate or comprehend utterances. This behavior is optimal because it minimizes the amount of time and mental energy spent in processing language, while still permitting successful communication—especially since people can typically monitor the success of communication through feedback from other interlocutors.

3. GROUNDING AND FEEDBACK

Language users require means for confirming understanding during communication (Clark, 1997; Clark & Brennan, 1991). Grounding refers to the process by which interlocutors collaborate to enter new information or a joint project into common ground. According to Clark and Brennan (1991), communicators present information to addressees and seek out positive evidence of that addressees have understood their intentions. Addressees, on the other hand, attempt to signal to communicators that they have accepted the information. Once both aspects of this process have taken place, a piece of information is effectively grounded.

The collaborative model assumes that people seek to ground information immediately and with a minimum expenditure of effort. The principle of opportunistic closure states that people want to ground their actions at the earliest available opportunity (Clark, 1997). They expect evidence that their goals, or joint projects, have been understood (and uptaken) before they move on to other things. Not only should the evidence be immediate, but the process of grounding should be accomplished in a way that minimizes collaborative effort (Clark & Wilkes-Gibbs, 1986). Thus, interlocutors will prefer presentations and uptakes which are unobtrusive, backgrounded, and require little work to interpret.

The idea of grounding has important implications for system design, because it provides guidelines for what kinds of information users will find helpful and when that information should be presented. If people obey the principle of opportunistic closure, then they should expect immediate and conspicuous evidence that their actions or messages have been understood. Many users are familiar with a common design problem in which the computer fails to signal

the uptake of a user-initiated action, such as when it carries out a time-consuming computation. In the absence of an explicit uptake, users will commonly repeat attempt to initiate the action over and over until they receive some sort of acknowledgment. Often these uptakes are delayed, or are not sustained over the course of the process. Another possible design problem concerns uptakes which are not backgrounded. To minimize effort and avoid disrupting user-interaction, a signal of uptake should be presented unobtrusively. Foregrounded and superfluous uptake signals such as needless dialog boxes and progress meters, might disrupt the flow of interaction. These could be replaced with more peripheral signals, like a sound which signals the execution of a command. Interestingly, the collaborative modelÕs emphasis on grounding appears to be at odds with the notion of optimal design. In a world in which communicators optimally design utterances for their addressees, and addressees exclusively use common ground as the context in which they comprehend utterances, why should language users require positive evidence of understanding? However, the fact that people need evidence of understanding fits very well with the picture of language use offered by the perspective adjustment model. Given that people typically speak and comprehend egocentrically, external feedback can trigger and facilitate the correction process if an error occurs. Speakers can gain knowledge about their addresseeÕs perspective with a minimum of effort by speaking egocentrically and monitoring feedback signals from their addressee.

4. SIGNALING

Because interlocutors seek to achieve closure on an action opportunistically, and to minimize their effort in doing so, it is important that the business of grounding not interfere with the business of talking. This is one reason why communication can unfold most naturally when the communicative medium allows them to take advantage of the three principal methods of signaling: describing-as, demonstrating, and indicating (Clark, 1997). When we think of language use, we typically think of the "describing-as" method, which is the use of abstract, conventionalized symbols to describe things. These signals are emitted primarily through the verbal channel. However, demonstrating and indicating, which often involve nonverbal displays, also play an important role. Sometimes demonstrating an action is more effective than describing it, such as when we want to teach someone how to change the tire on a car. Furthermore, by indicating objects or situations nonverbally (e.g., by pointing or looking at them), people can call othersÕ attention to them without having to describe them. In fact, pointing to a referent object while speaking can sometimes change the meaning of what is said (Barr & Kelly, 1997).

Traditionally, system designers have focused on enabling the "describing-as" method of signaling. For example, a software designer selects a combination of keystrokes to symbolize a certain action, or the inventor of a programming language chooses a string of symbols to represent a sequence of operations. When designers develop systems to allow communication between humans, they typically focus on enabling interchange through a written or spoken channel, which impairs the ability of communicators to indicate objects or demonstrate actions, because these methods of signaling typically involve a large nonverbal component. Although symbolic interchange in the computer medium is important, in the real world signals between interlocutors typically are a multi-modal composite of describing, indicating, and demonstrating (Clark, 1997). Requiring users to squeeze information through the symbolic bottleneck can make communication laborious and error-prone. For example, users and support technicians often encounter difficulty and confusion when trying to diagnose a computer problem over the telephone. In many cases, a user would find it easier and more natural to visually demonstrate a problem to technicians than to attempt to describe it to them.

In addition to these design considerations, the theory of signaling provides us with a means of evaluating the appropriateness of various communications technologies for different cooperative tasks. People can communicate in many different ways: by teleconferencing, fax, email, or telephone, among many others. Communication media vary in the degree to which they enable people to describe, indicate, or demonstrate. By choosing a medium which is most suited to the communicative requirements of a task, we might be able to increase the chances that cooperation will be successful.

ClarkÕs work on signaling represents an important contribution to research on language use. However, these claims need to be empirically evaluated. Moreover, we need to gain a better understanding of how these different kinds of signals function in language production and perception. For example: How do they interact to create meaning, both when people formulate and understand utterances? What happens when one or more of these methods is impaired? How sensitive are comprehenders to different kinds of signals?

5. SOME PRINCIPLES FOR DESIGNING COOPERATIVE SYSTEMS

To summarize, we have identified the following application areas of the collaborative model to the design of cooperative computer systems: reference diaries, optimal design, grounding, and signaling. In each of these areas, the collaborative model provides potentially useful guidelines for designing systems which support language use. However, we would like to emphasize the need to empirically evaluate the claims made by the collaborative model. Many claims of the model are without empirical support. Some studies which appear to provide evidence for the use of common ground in comprehension do not in fact do so, because of a design confound (Keysar, 1997). Empirical research which unconfounds these factors shows that even when language users have detailed reference diaries they donÕt use them for Òoptimal design.Ó This state of affairs might suggest quite a distinct Ôtheory of the userÕ from that proposed by the collaborative model.

In order to design systems which feel natural to users, are easy to learn how to use, and enable efficient and successful communication, designers might consider the following proposals about how people process language:

People appear to be systematically egocentric in interpreting and formulating messages (Keysar, Barr & Horton, in press). The perspective adjustment model suggests that common ground is used primarily in error correction. According to the collaborative model, people seek positive evidence that their actions have been understood. They expect that grounding signals will occur immediately and in an unobtrusive manner.

The model suggests that people communicate most naturally when the medium of communication allows them to indicate and demonstrate as well as describe. To date, most systems have been designed principally to enable describing, but typically neglect people's need to indicate objects to one another and to demonstrate actions.

6. ACKNOWLEDGMENTS

The writing of this paper was facilitated by the generous support of a John Simon Guggenheim Memorial Fellowship and a PHS grant R29 MH49685 to Boaz Keysar.

7. **REFERENCES**

Barr, D. J. & Kelly, S. D. (1997) Nonverbal factors in understanding and remembering indirect requests. In M. Shafto & P. Langley (Eds.), Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society. Hillsdale, N. J. : Lawrence Erlbaum.

Clark, H. H. (1997) Using language. Cambridge : Cambridge University Press.

Clark, H. H. & Brennan, S. E. (1991) Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), Perspectives on Socially Shared Cognition (pp. 127-149). Washington, D. C. : American Psychological Association.

Clark, H. H. & Carlson, T. B. (1981) Context for comprehension. In J. Long & A. Baddeley (Eds.), Attention and Performance IX (pp. 313-330). Hillsdale, N. J. : Lawrence Erlbaum.

Clark, H. H. & Marshall, C. R. (1981) Definite reference and mutual knowledge. In A. H. Joshe, B. Webber & I. A. Sag (Eds.), Elements of discourse understanding (pp. 10-63). Cambridge : Cambridge University Press.

Clark, H. H., Schreuder, R. & Buttrick, S. (1983) Common ground and the understanding of demonstrative reference. Journal of Verbal Learning and Verbal Behavior, 22, 245-258.

Hanna, J. E., Trueswell, J. C., Tanenhaus, M. K. & Novick, J. M. (1997) Consulting common ground during referential interpretation. Poster presented at the 38th Annual Meeting of the Psychonomic Society. Philadelphia, PA.

Horton, W. S. & Keysar, B. (1996) When do speakers take into account common ground? Cognition, 59, 97-117. Johnson-Laird, P. N. (1982) Mutual ignorance: Comments on Clark and CarlsonÕs paper. In N. V. Smith (Ed.), Mutual Knowledge. London: Academic Press.

Keysar, B. (1997) Unconfounding common ground. Discourse Processes, 24, 253-270.

Keysar, B. (1994) The illusory transparency of intention: Linguistic perspective taking in text. Cognitive Psychology, 26, 165-208.

Keysar, B., Barr, D. J., Balin, J. A. & Paek, T. (in press) Definite reference and mutual knowledge: Process models of common ground in comprehension. Journal of Memory and Language.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (1996) Common ground: An error-correction mechanism in comprehension. Paper presented at the 37th Annual Meeting of the Psychonomic Society. Chicago, IL.

Sperber, D. & Wilson, D. (1982) Mutual knowledge and relevance in theories of comprehension. In N. Smith (Ed.), Mutual Knowledge. London: Academic Press.