Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Interaction Organization in Graphical Communication

Permalink

https://escholarship.org/uc/item/5h3883v5

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 25(25)

ISSN

1069-7977

Authors

Umata, Ichiro
Shimojima, Atsushi
Katagiri, Yasuhiro
et al.

Publication Date

2003

Peer reviewed

Interaction Organization in Graphical Communication

Ichiro Umata
umata@atr.co.jp
ATR Media Information Science Laboratories;
Seika Soraku Kyoto, 619-0288 Japan

Atsushi Shimojima
ashimoji@jaist.ac.jp
Japan Advanced Institute of Science and Technology;
1–1 Asahi Tatsunokuchi Nomi Ishikawa 923-1292 Japan

Yasuhiro Katagiri **Nik Swoboda**
{katagiri/nswoboda}@atr.co.jp
ATR Media Information Science Laboratories;
Seika Soraku Kyoto, 619-0288 Japan

Abstract

In this paper, we examine interaction organizations in communication through drawing. We report that both sequential and parallel interaction organizations take place in actual human graphical communication, despite the fact that the media use involved in drawing has resource characteristics that better match parallel organizations. We argue that activity-dependent constraints imposed by the task performed in the interaction play a significant role in determining the interaction organization. We propose two types of sequentiality constraints and demonstrate that they can override the resource characteristics of the drawing media, thereby enforcing a sequential turn-taking organization similar to those observed in verbal interactions.

Introduction

Turn taking and simultaneous action are two contrasting organizational principles for multi-party joint activities. People take turns to speak in conversation, whereas musicians play their instruments simultaneously in orchestra. One possible explanation for this contrast is that the resource characteristics of the media use involved in the activity dictates the style of interaction. The media use for speech affords only one person's speech sounds at a time, whereas the media use for music affords multiple instrument sounds simultaneously.

Sacks et al. (1974) regard verbal turns as an economic resource, which are distributed to conversation participants according to turn organization rules. According to them, one of the main effects of these turn organization rules is the sequentiality of utterances, namely, the almost complete absence of parallel or simultaneous utterances from conversation. "Overwhelmingly, one party talks at a time," they observe. On the other hand, Brennan (1990) reported that non-verbal signals, such as the movement of a cursor, did not observe the turn organization rules. Condon (1971) even claims that the gestures and other bodily movements of participants in conversation exhibit mutual synchrony with each other. These findings apparently suggest that verbal and non-verbal media in interaction do not necessarily adhere to the same turn organization rules and that these differences may derive from the contrasting resource characteristics of the media used in the interaction.

The first purpose of this paper is to investigate whether the turn organization holds for drawing interaction. People engaged in spoken conversation often draw letters, text, figures, and pictures on a piece of paper, a chalkboard, or whatever affords a drawing surface. Do such drawing expressions also obey certain turn organization rules? In particular, are they organized sequentially, without substantial temporal overlapping? Or do people freely draw in parallel?

The second purpose is to elucidate the dominant factors in deciding between the two interaction organizations for drawing. Exactly when is sequential turn taking for drawing mandated? When is simultaneous drawing possible? [1]

Although theoretical in nature, these questions have been mainly addressed in HCI literature, receiving conflicting answers. Stefik et al. (1987) report on a computational facility where multiple people engaged in spoken dialogue can simultaneously "draw" on a shared video display. The design idea was that this functionality would liberate drawing expression from the sequentiality constraint typical in spoken expressions, giving everyone more chances to express herself. Thus, a key assumption of their work is that people would use simultaneous drawing for communication, given a facility to do so.

This assumption was also partially supported by Whittakker et al. (1991), who observed six groups of subjects using a virtual shared whiteboard for long-distance communication. They found frequent instances of simultaneous drawing on the whiteboard, especially, when diagrams and tables were drawn or modified.

Tatar et al. (1991), however, give a contradictory picture in their observational study of two groups of subjects using the same system that Stefik et al. (1987) described. They found that despite the potential for simultaneous drawing, the subject interactions were largely sequential, trying to take drawing turns as well as speaking turns. They explained this fact by claiming that human communication generally proceeds incrementally, grounding expressed information one by one in real time. According to them, the facility for simultaneous drawing is based on an incorrect model of human communication, redundant at best and troublesome at worst.

We argue, in this paper, that the resource characteris-

---

[1] Here by *simultaneous drawing* we mean drawing activities in which more than one participant is drawing at the same time. This phenomenon excludes instances when simultaneous drawing and erasing or erasing and erasing occur.

tics of the media involved are not sufficient to provide a full account of the interaction organization, and that constraints imposed by the task performed in the interaction also play a significant role. We propose two types of task constraints and show how they work in organizing interaction through the examination of data we collected in experiments on human communication by drawing over electronic shared whiteboards.

Hypothesis

One may attempt to understand the reasons for the sequentiality of speaking turns by citing two factors. The first is the *non-persistence* of speech. Speech occurs, passes away, and leaves little trace once it is over. Speech therefore must be comprehended in real time, and its content must be grounded without too much delay, usually before the next substantial utterances are made. Grounding of verbal utterances must be thus *incremental*.[2]

The second factor is the *restricted bandwidth* of speech as a communication medium. Because of the limitations of human auditory processing, two or more utterances made at the same time are difficult to understand separately. Thus, it is unusual for more than one message to be simultaneously transmitted through speech, except when they are acoustically similar to make so-called “sync talk”.

Given this non-persistence and bandwidth limitation, it is very difficult to ground two spoken utterances made at the same time. You cannot comprehend them together because of the bandwidth limitation, while you cannot delay the comprehension of one utterance until later, because of its non-persistent character. Thus, it is preferable that verbal utterances be organized in sequential turns, without substantial temporal overlap.

If one follows this explanation, one naturally thinks of cases where such sequential turns are *not* mandatory. Drawing is one interesting example of such a case. First, it is usually persistent. The drawing activity itself occurs and passes away just as speech does, but it leaves a trace on the drawing surface. Thus, the grounding of drawing content does not have to be made in real time, and can be delayed until later. Second, with a sufficiently large drawing surface, two or more drawing operations can occur at the same time, without the fear of their content being mixed in an incomprehensible manner. Thus, drawing has larger bandwidth than speech does. For these reasons, one may expect that drawing can occur in parallel, without sacrificing natural communication. Simultaneous drawing may be even preferred, since it enables the expression of a larger amount of information in a given time than sequential drawing does.

This theory endorses the assumption underlying the communication facility reported by Stefik et al. (1978), that people use simultaneous drawing when possible. It also explains the frequent occurrences of simultaneous inputs observed by Whittaker et al. (1991). The theory, however, sharply contradicts the finding and claim by Tatar et al. (1991).

Let us therefore call this theory *Parallelism*. It claims that drawing expressions in communication have no mandatory sequential structure. In contrast, *Sequentialism* is the view that sequentiality is a requirement on drawing expressions.

In this paper, we will posit and test yet another view as our hypothesis. According to this view, sequential structure is mandatory in drawing in some cases but not in others. More specifically:

Sequentiality Constraint Hypothesis

1. Drawing interaction occurs in sequential turns under either of the following conditions:
   (a) Information Dependency Condition: when there is a dependency among the information to be expressed by drawing,
   (b) Event Alignment Condition: when drawing operations themselves are used as expressions of the proceedings of target events.
2. Sequential turns are not mandatory in drawing activities when neither condition holds (and persistence and certain bandwidth of drawing are provided.)

The rationale for the information dependency condition is the intuition that when one piece of information depends on another, the grounding of the former piece of information is more efficient *after* the grounding of the latter has been completed. This should be the case whether a particular speaker is explaining the logical dependency in question to her partners, or all participants are following the logical steps together.

For example, suppose you and your partner are collaboratively drawing a map of a town on the basis of your memory. Figure 1 is the map that is currently on their sheet of paper. Given this map, the information (3) is a logical consequence of the information (1) and (2):

(1) The Bank is in front of the Post Office across Seventh Avenue.

(2) The City Hall is in front of the Bank across Maple Street.

(3) The building represented by the blank square is the City Hall.

Our point is that when you are grounding these pieces of information, you will do so incrementally, starting from (1) and (2), and then moving onto (3). The dependency of the information encourages or even enforces incremental grounding.

Note that once incremental grounding is employed, interactions take place solely on a local area of the drawing surface where the relevant information is expressed. The size of the drawing surface therefore becomes irrelevant, and the bandwidth afforded by it remains unused.

In this case, because the persistency and bandwidth of drawing is not being effectively used, drawing is in

---

[2] See Clark & Schaefer (1989), Clark (1996), and Traum (1994) for the notion of grounding. Chapter 8 of Clark (1996) contains a detailed discussion of incremental grounding.

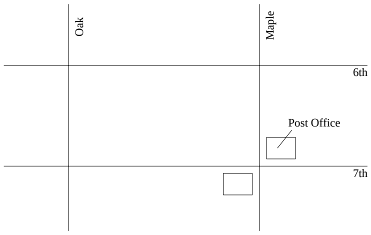

Figure 1: A townmap on which information (1), (2), and (3) will be drawn.

the same situation as speech. Sequential turn structure is mandatory, generating sequential drawing turns. In the above example, you would draw an icon labeled “Bank” first, ground the expressed information, and then label the blank square as “City Hall”.

Event alignment is a strategy for expressing the unfolding of an event dynamically, using the process of drawing itself as a representation. For example, suppose you are reporting how you spent a day in the town depicted in Figure 1. You might say that you walked from Oak and Seventh to Maple and Sixth via Maple and Seventh. If you report this process while drawing a line along the route in the map, saying “I started from here and turned left at here...”, you are using the drawing process as a representation of your walking process. You are aligning the drawing event with the walking event to express the latter dynamically.

Our hypothesis is that simultaneous drawing is unlikely while this strategy of event alignment is employed. Under this condition, the movement or process of drawing is the main carrier of information. The trace of drawing has only subsidiary informational roles. Thus, in this particular use of drawing, its persistency is largely irrelevant. The message must be comprehended and grounded in real time, and bandwidth afforded by the drawing surface becomes also irrelevant. This requirement effectively prohibits the occurrence of any other simultaneous drawing.

Method

An experiment in which subjects were asked to communicate graphically was conducted to examine the effect of the two factors presented above on their drawing activities. In these experiments, 24 pairs of subjects were asked to work collaboratively on four problem solving tasks using virtual whiteboards.

Experimental Setting

In the experiments being reported here, two subjects collaboratively worked on four different problem solving tasks. All the subjects were recruited from local universities and were paid a small honorarium for their participation. The subjects were seated in separate, soundproof rooms and worked together as a pair using a shared virtual whiteboard (50 inches) and a full duplex audio connection. The subjects were video-taped during the experiment. They also wore a cap-like eye-tracking device which provided data indicating their the eye-gaze positions. The order in which the tasks were presented was balanced between the 24 pairs so that the presentation order would not have an affect on the results. The time limit for each task was six minutes.

At the start of each task, an initial diagram was shown on both of the subject’s whiteboards and then subjects were free to speak to one another and to draw and erase on the whiteboard. The only limitation to this drawing activity was that they could not erase or occlude the initial diagram. All drawing activity on the whiteboard was performed with a hand-held stylus directly onto the screen, and any writing or erasing by one participant appeared simultaneously on their partner’s whiteboard. The stylus controlled the position of the mouse pointer, and when not drawing, the position of both subject’s mouse pointers was displayed on their partner’s whiteboard.

Tasks

Deduction Task with an Event Answer (1e) A logical reasoning problem with a correct answer. The problem asks that the subjects describe the arrangement of people around a table and the order in which people sit down. This seating arrangement and order must satisfy some restrictions (e.g. “The fifth person to sit is located on the left-hand side of person B”). A circle representing a round table was shown on the whiteboards at the start of the task. This task has strong informational dependency and strong event alignment.

Deduction Task with a State Answer (1s) A logical reasoning problem with a correct answer asking that the subjects design a seating arrangement satisfying some restrictions (e.g. “S cannot sit next to M”). A circle representing a round table was shown on the whiteboards at the start of the task. This task has strong informational dependency and loose event alignment.

Design Task with an Event Answer (2e) A task with an open-ended answer, asking subjects to make an excursion itinerary based on a given town map. A complete town map was shown on the whiteboards at the start of the task. This task has weak informational dependency and strong event alignment.

Design Task with a State Answer (2s) A task with an open-ended answer, asking the subjects to design a town layout to their own likings. An incomplete town map was shown on the whiteboards at the start of the task. This task has weak informational dependency and loose event alignment.

Data

During each task all drawing, erasing, and mouse movements by each subject was recorded in a data file. Using this data, the amount of simultaneous drawing was calculated as the percentage of the total time spent drawing simultaneously of the total time either subject spent drawing (i.e. the sum of the time intervals in which both

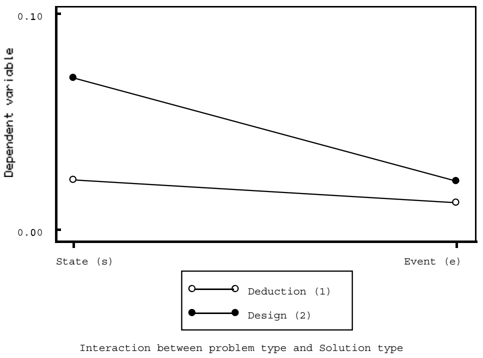

Figure 2: Amount of simultaneous drawings

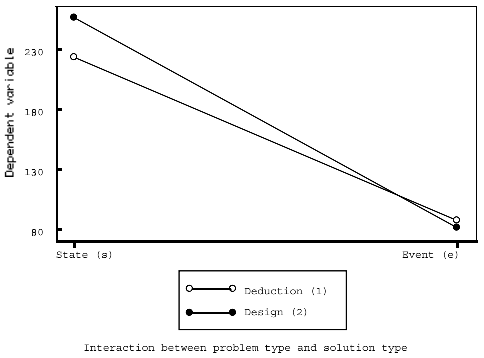

Figure 3: Average distance of simultaneous drawings

subjects drew simultaneously divided by the the sum of the time intervals in which at least one of the pairs drew on the whiteboard). The average distance between the pens during each simultaneous drawing event was calculated based upon the averaging of the distances in pixels between the two pens sampled every 200 msec during all simultaneous drawing events.

Results

The Amount of Simultaneous Drawing

As is shown in Figure 2, the proportion of simultaneous drawing time to total drawing time was the largest in the design state (2s) condition. This data was entered into a 2 x 2 Analysis of Variance (ANOVA). Both problem type (deduction and design) and by solution type (state and event) were treated as within subject factors. Analyses revealed a main effect of problem type F(1,23)=15.25, p<.001 and solution type F(1,23)=32.033, p<.001 which was qualified by an interaction F(1,23)=8.33, p<.01. This interaction was caused by an effect of solution type on the design task F(1,23)=42.88, p<.001. No effects were found for the solution type on the deduction task F(1,23)=2.13, p<1 nor for the problem type on the event solution F<1.

Thus, it was shown that simultaneous drawing tends to be blocked when the task has either strong informational dependency or tight event alignment, or both of them.

The Average Distance of Simultaneous Drawings

Figure 3 illustrates the average distances between the pens during simultaneous drawing. It can be seen that the event tasks had shorter average distance than the state tasks. This data was again entered into the same 2 x 2 Analysis of Variance (ANOVA) used in the last section. Analyses showed a simple main effect of solution type F(1,23)=23.19, p<.001 and no effect of problem type nor interaction Fs<1.

Thus, it was found that simultaneous drawing happens closer to each other when the task has strong event alignment.

Evidence for the sequentiality of constraints

The sequentiality of constraints manifest themselves in the actual behavior of participants in graphical communication. We found evidence in the recorded data of interactions that indicate the effects of the constraints.

Information dependency

Figure 4 shows a snapshot of the drawing interaction that took place between two collaborators working on the design state task (2s). The upper row and the lower row indicate the drawing behaviors of subjects *A* and *B*, respectively. No information dependency was imposed by the task, and the event alignment strategy was neither necessary nor plausible. *A* and *B* agreed to divide the design task into two sub-tasks, the design of a station plaza and the design of a park, and started to work on each separately. *A* is now drawing a forest icon, while *B* is drawing a station sign simultaneously. They declare what they intend to draw when they start drawing, but they do not acknowledge the drawing produced by their partners until they finish their own drawings. This example indicates that, without the sequentiality constraints, collaborators are free to adopt rather coarse-grained incremental grounding activity style, which in turn permits parallel drawing.

We observed quite different behaviors in the deduction state task (1s). Information dependency is created here by the logical dependency between seating assignments. Two subjects *A* and *B* are discussing at one point during the course of their interaction which seat to assign to a particular person *S* at a round table before moving on to work on other people. After enumerating relevant conditions posed in the problem, they start agreeing on one particular slot. While they are both verbally referring to the position of the slot, *A* starts drawing the mark S in the slot to indicate the placement of *S*. *B*, who also started to initiate the same drawing action, stops in the middle when he observed *A*'s drawing action, and verbally acknowledges *A*'s action instead. This example clearly shows that the information dependency constraint posed in the task brings about fine-grained incremental grounding activity between collaborators by making them work in close agreement regarding the state of their partial so-

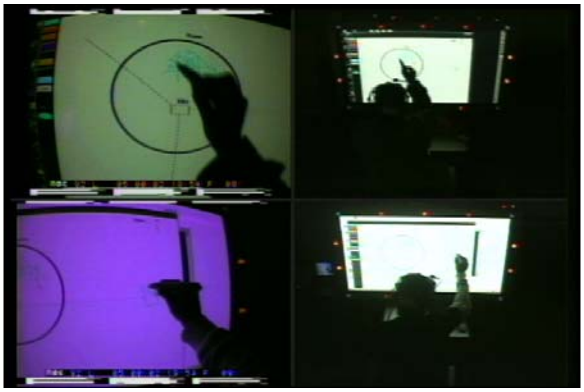

Figure 4: Drawing interaction exhibiting coarse-grained incremental grounding in the design state condition.

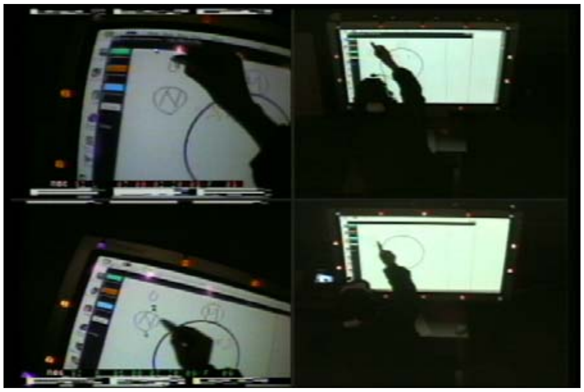

Figure 5: Drawing interaction without the event alignment strategy in the deduction state condition.

lution and aligns decisions about which seating slot to assign to whom. This activity style, then, forces them to observe sequential turn-taking organization and to avoid simultaneous drawing.

Event alignment

The event alignment strategy was evident in the design event task (2e). A typical case was observed in which subjects *A* and *B*, while jointly designing a tour in the virtual town, have agreed to visit the museum next and have just begun discussing the choice of transportation means. *A* first proposes to go on foot. *B* makes a counterproposal to go by a train. *A* agrees. *A* and *B* both start to initiate the drawing of the line for the move on the map. *B* is the first to actually initiate the drawing. *B* draws the line to the museum icon while simultaneously describing the movement by the utterance "go like this." *A*, upon looking at *B*'s drawing already initiated, pulls his hand away from the whiteboard, and produces verbal acknowledgment to *B*'s drawing action. *B*'s drawing here embodies *B*'s adoption of the event alignment strategy. The event of moving the pen to the museum icon represents the event of being transported by train to the museum in the designed tour. The event alignment strategy enforces the sequential organization of the interaction. Evidence for this hypothesis is given in the example by the fact that they avoided jumbled drawing and restrict themselves to turn taking in their drawing productions.

Figure 5 shows a snapshot from the deductive state task (1s). The logical dependency between seating assignments works as the information dependency constraint, but the event alignment strategy does not occur here. Subjects *A* and *B* are discussing who to seat next to *M*. *B* draws the sign N as one candidate, while *A* simultaneously draws the sign O, saying "*O* is also possible." Simultaneous drawing occurs for disjunctive pieces of information. This example suggests that when the event alignment strategy is not adopted, the information dependency constraint alone leaves room for parallelism.

Even though information dependency constraints enforce incremental grounding interaction styles, it is still possible to find room for parallelism within each incremental step, as long as the pieces of information expressed during simultaneous drawing are not subject to any information dependency constraints. On the contrary, the event alignment strategy enforces close coordination of attention to the same object at the same time between collaborators. This requirement for coordination works strongly to favor the sequential organization of the interaction, because it effectively prevents collaborators from exploiting both the persistence and the bandwidth of the drawing media.

Discussions

Sequentiality hypothesis and drawing turns

Our sequentiality hypothesis predicts that even with persistent media, e.g., a whiteboard, collaborative activities would exhibit sequential turn-taking organizations when the joint task conducted using the media demands either or both of the information dependency and the event alignment constraints be satisfied. Our experimental results shown in Figure 2 supports the prediction. We observed frequent occurrences of simultaneous drawing in the design state task condition. About one sixth of one's entire drawings were made simultaneously with those by her partner. But, significantly less temporal overlap in drawing were observed in the other three task conditions: design event, deduction state and deduction event. Simultaneous drawing is possible when neither of the two sequentiality constraints are imposed by the task, whereas turn-taking behaviors are enforced by either of the sequentiality constraints.

Simultaneous drawing distance

Even though both the information dependency and the event alignment constraints work equally in enforcing the sequential turn-taking organization, these two factors manifest themselves in different drawing behaviors when two participants draw simultaneously, as was observed in Figure 3. Simultaneous drawing produced in the state conditions was placed, on average, farther away from each other than those produced in the event conditions. It appears the event alignment strategy, which is a good match for working in the event conditions, is the major factor in determining how close simultaneous drawing can take place. This is understandable when we consider that the event alignment strategy effectively suppresses the bandwidth characteristics of drawing media by forcing collaborators to focus on the same region of the whiteboard. Simultaneous drawing must be the accidental crash of drawings, similar to the crash of utterances observed in verbal interactions. As long as the event alignment strategy is not adopted, the information dependency constraint alone, which is effective in the logic state condition, still leaves room for collaborative simultaneous drawing, as we saw in the example in Figure 5, which can be produced in farther away positions.

Parallelists vs. Sequentialists

How do our results compare with claims by Parallelists and Sequentialists? Parallelists claim that the persistence of media eliminates the requirement of incremental grounding in interaction, making it possible for interactants to produce drawings in parallel. However, even though persistent media makes it possible, in principle, for collaborators to adopt non-incremental grounding strategies, there can be other sources, e.g, the sequentiality constraints, that demand incrementality in grounding. When the task itself demands the participants to observe certain information dependencies, either logical or temporal, the collaborators have to design their activities so that they attain the alignment of knowledge and the agreement of their decisions before they move on to the next step. Otherwise their activities diverge from each other so that they quickly lose track of dependency relationships between pieces of information and run the risk of unsuccessfully completing the task.

On the other hand, Sequentialists criticize the parallel use of persistent drawing media in interaction on the grounds that natural interaction needs to be sequential. Our design state task condition, as well as other studies on parallel drawing activities, clearly show that natural parallel drawing interactions are possible. This is because where neither the information dependency constraint nor the event alignment strategy is effective, collaborators can fully make use of the persistence of the drawing media to separate the production and the comprehension of drawings, thereby making one's production of drawings independent from that of her partner.

In short, the characteristics of the media is not the only determinant of the turn-taking organization in interaction. The constraints in the task and the strategies adopted for the activities both play significant roles in what kind of organization structures emerge in collaborative interactions.

Conclusions

We examined how interactions are organized in graphical communication. Following the mixed results obtained in the HCI field, we first postulated two types of constraints that mandate sequential drawing turns. We conducted an experimental analysis of actual human graphical interactions over an electronic shared whiteboard, putting them under the four conditions corresponding to the presence and the absence of the postulated constraints. Differences were observed both in terms of the amount of and the distance between simultaneous drawings, which confirmed the effectiveness of the constraints in determining sequential/parallel interaction organizations. A detailed examination of the data also revealed additional evidence for the effectiveness of the constraints. These results lead us to a new picture of interaction organization mechanism in which both the activity-level constraints imposed by the task and the resource-level constraints imposed by the media's physical characteristics affect the interactional styles in communication. These findings suggest that task-level constraints need to be taken into consideration when designing interactions in HCI and CSCW systems. Exactly how they should be considered, of course, depends on further elucidation of the task-level constraints and the roles they play in various types of interactions.

Acknowledgments

This research was supported in part by the Telecommunications Advancement Organization of Japan.

References

Brennan, S. E. (1990). Seeking and providing evidence for mutual understanding. Ph.D. dissertation, Stanford University.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H. and Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13*, 259–294.

Condon, W. S. (1971). Speech and body motion synchrony of the speaker-hearer. In D. L. Horton and J. J. Jenkins (Eds.), *Perception of Language*, Columbus, Ohio: Merrill, 150–173.

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

Stefik, M., Foster, G., Bobrow, D., Kahn, K., Lanning, S., and Suchman, L., (1987). Beyond the Chalkboard: Computer Support for Collaboration and Problem Solving in Meetings. *Communications of the ACM, 30 (1)*, 32–47.

Tatar, D., Foster, G., and Bobrow, D. (1991). Design for Conversation: Lessons from Cognoter, *International Journal of Man-Machine Studies, 34(2)*, 185–210.

Traum, D. (1994). A computational theory of grounding in natural language conversation. Ph.D. dissertation, University of Rochester.

Whittaker, S., Brennan, S., and Clark, H. (1991). Coordinating activity: an analysis of computer supported co-operative work. In *Proceedings of CHI' 91 Human Factors in Computing Systems*, 361–367, New York: ACM Press.