Explicit Referencing in Chat Supports Collaborative Learning

Martin Mühlpfordt
Fraunhofer Institute for Integrated Publication and Information Systems (IPSI)
{martin.muehlpfordt, martin.wessner}@ipsi.fraunhofer.de

Martin Wessner

Abstract. In Pfister & Mühlpfordt (2002) a study was presented showing that chat discussions with a strict turn order combined with the requirement to assign a type and an explicit reference to each message lead to a higher learning score than discussions in a normal chat or in a chat with strict turn order only. Due to the experimental design it was not possible to judge the role of explicit referencing. Now we present the "missing" data: The higher learning score can be explained just by the explicit referencing. We argue that this is an important design issue for chat applications, because it seems that explicit referencing leads to a more homogeneous discourse behavior (more homogeneous participation, more participation in parallel discussion threads) and a better grounding. A case study explored the use of the referencing function in a less restricted everyday collaborative situation.

Keywords: CSCL, Chat, Referencing, Experimental Study, Case Study

INTRODUCTION

Collaborative learning relies on successful communication. Successful means that the collaborators understand each other's contributions and build a shared understanding of the collaboration content. When the group communicates using chat, the communication is influenced by the medial properties. A prominent phenomenon of chat communication is the somewhat "chaotic" discourse structure: Often the group discusses two or more topics in parallel, and related turns are in contrast to spoken conversations not adjacent. Another important requirement for effective collaborative learning is the combination of communication with shared artifacts (artifact centered discourse; Suthers & Xu, 2002).

In this paper, we propose that extending the medium chat with the possibility to assign explicit references to a message and to shared material supports the group in their discourse.

The remainder of the paper is structured as follows. In the next section, grounding and discourse comprehension are presented as a theoretical basis for explicit referencing. Then we present an empirical study, which shows that explicit referencing results in a higher learning score. A post-hoc analysis of the discourses indicates that grounding is supported by explicit references (section 3). To explore acceptance and usage of explicit references in less restricted chats, we conducted a case study in an everyday collaborative situation, which is presented in section 4. In section 5 we compare our approach to related work. Finally, we conclude with a discussion of the results and identify open research questions.

THEORETICAL BACKGROUND

In this section, we present some theoretical background for the explicit referencing. We argue that explicit referencing influences the grounding strategies of the communication partners and eases the process of discourse comprehension. Therefore, we shortly present the concept of grounding and theories of discourse comprehension. Then we discuss findings of chat research in respect to these aspects. Finally we describe the concept of explicit referencing in more detail.

Grounding

Building a common ground is an integral part of collaborative learning (Baker et al., 1999). The communication partners construct a shared understanding of what is said by giving mutual feedback. They reciprocally ensure each other that the ongoing discourse builds on and extends shared knowledge. Only the contributions that are grounded (mutually believed to be understood) become part of that shared knowledge, the so-called common ground (Clark, 1996). For achieving and maintaining the common ground in spoken face-to-face communication, various forms of linguistic and non-linguistic feedback are used. People use methods like giving oral feedback (e.g. "hm"), non-verbal attention cues (e.g. eye contact), or initiating turn-taking (e.g.
asking a question) for grounding. Mostly, positive feedback that an utterance is understood is given simultaneously without interrupting the speaker’s turn.

The effort for grounding required by the participants varies with the properties of the communication medium (Clark & Brennan, 1991). The medium constrains the communication and thereby influences the costs of grounding. For instance, the medium may constrain the people with respect to cotemporality (Can the production of an utterance be perceived by the communication partners roughly at the same time?) or sequentiality (Can the turns get out of sequence?).

Following the principle of least collaborative effort, that “in a conversation the participants try to minimize their collaborative effort – the work that both do from initiation of each contribution to its mutual acceptance” (Clark & Brennan, 1991, p. 135), different media result in different styles of grounding.

Discourse comprehension

The concept of grounding describes the mechanisms used by the communication partners to ensure and keep track of the mutual understanding. Theories about discourse comprehension try to explain how people comprehend, what others say. An important aspect of the comprehension process is to infer the relation – the local coherence – of the utterance to the surrounding discourse (Hobbs, 1985). Elliptic and anaphoric expressions refer back to objects introduced earlier in the discourse, and descriptive referential expressions point directly to a previous utterance or paraphrase it (Eklundh & Rodriguez, 2004). This inference process can be treated as problem solving to arrive at the speaker’s intended interpretation (Clark, 1978).

Chat research

Chat is widely used to do conversations online. From the linguistic perspective chat-conversations share some features of oral language although it is a textual medium (Koch & Österreicher, 1994), like accepting surface errors (syntax and grammar), using informal phrases etc. It is claimed that this results from the similarities between the communication situations (Murray, 2000), more specifically from the communicative attitude in a spoken face-to-face communication and a chat conversation. Nevertheless, the medial properties of chat lead to discourse structures that are different to the ones of from spoken face-to-face conversations. These medial properties are:

1. Separation of production and presentation: The production of an utterance (chat message) cannot be perceived by the communication partners. The message is presented to them as a whole only after it has been sent by the contributor.
2. Sequencing of messages: The contributor cannot determine the exact position of a message, which depends on the simultaneous (not observable) communication behavior of the communication partners.

As a consequence, the sequential order of the messages is characterized by disrupted turn adjacency (Herring, 1999). This complicates the identification of the message, the new one is responding to and leads to so-called phantom adjacency pairs, i.e. pairs of subsequent messages that seem to be related, but are not intentionally related (Garcia & Jacobs, 1998). Another consequence is the lack on simultaneous feedback, which is especially important for the grounding strategies in spoken communication.

Facing these problems chat users adapt their communication strategies: For instance (1) a turn is broken down in a sequence of messages, indicating at the end of the message, that it is going to be continued (Herring, 1999). (2) The addressee of the message is explicitly mentioned. (3) Responses to messages of parallel ongoing threads are done with different messages, preserving the "inner-thread" sequential order (O'Neill & Martin, 2003). This shifts at least partially the costs from understanding to production of a message.

Concept of Explicit Referencing

Explicit referencing means that while producing a message an object in the shared environment can be selected and assigned as a referential point to that message. We call the (directed) relation between the message and this referential point a reference. This reference is transmitted together with the chat message and visualized as an arrow starting at the message and pointing to the referential point. A previous message, a portion of it, or some part of a material viewed in the shared environment can be selected as referential point. When receiving a message, the reference is automatically shown. The reference of a previous message can be made visible by selecting that message (see reference from a contribution to a part of the shared material in fig. 1).
We expect that explicit referencing influences the communication in the following ways:

1. Message comprehension: The inference of a message’s relation to the surrounding discourse (including other messages and shared material) is simplified.

2. Message production: As the costs of understanding a non-adjacent message are reduced, participants are encouraged to respond also to older messages. This leads to more grounding activities, e.g., acknowledging previous messages. In addition, referencing to shared material saves production costs as effort to repeat or describe parts of the material is reduced.

EFFECT OF EXPLICIT REFERENCING TO LEARNING SUCCESS

In a prior study (Pfister, Mühlpfordt & Müller, 2003, preliminary results were presented in Pfister & Mühlpfordt, 2002), the effect of system controlled so-called learning protocols on learning was explored. In that study, the learning protocol controlled the turn order with only one participant contributing at a time.

In this study, learning under three different conditions was compared. (1) The complete protocol condition. Here, the participants were forced to create an explicit reference before actually writing the message, and the turn order was system controlled. (2) The chat application used for the second condition didn’t have the referencing function, but the strict turn order was applied. (3) The control condition. Here, the participants used a normal chat tool without referencing and without turn control.

Three different group sizes (dyads, triads, and quartets) and two different learning domains (causes and consequences and different types of earthquakes vs. difference between to opine, to belief and to know) were tested. The study showed for the earthquake domain superior learning across the different group sizes under the complete protocol condition, and no effect of the strict turn order without referencing. Because of the design of that study, it was not possible to check whether the superior learning was due to the explicit referencing alone or the combination of both, strict turn control and explicit referencing. Therefore we extended that study by the missing forth condition: learning without strict turn control and with explicit referencing.

Method

We decided to check only triads and quartets, because we expected that the impact of the explicit referencing is low for dyads as only one other participant can disrupt the sequential turn order. Furthermore, we focused on the earthquake domain, which previously has shown a clear effect. Together with the conditions of the prior study, we have a three factorial between-subjects design with the factors Turn-Control (with and without), Referencing (with and without), and Group Size (triads and quartets).

A total of 31 subjects (students of the Technical University of Darmstadt) participated in the study, put together in 5 groups of 3 and 4 groups of 4. Participants received 15 Euro for participation in a one-hour
The sessions were organized according to the prior study: The learners and the tutor worked with standard PCs in isolated cubicles (i.e. they could not see each other), simulating a distributed scenario. First, the general scenario was introduced, a short questionnaire about experience with computers and the internet was administered, and the user interface of the chat tool was explained. The participants learned how to refer to previous messages and the material and how the messages are actually written and sent. Then, a short knowledge test was applied to assess participants' degree of prior knowledge. The test consisted of one open question (“Explain shortly how earthquakes evolve.”). Then, the participants started with the learning process. A learning goal was provided and presented on a sheet of paper attached to the PC. The learning goal was “to understand causes and consequences of earthquakes and different types of earthquakes”. The time limit for a learning session was 25 minutes. The tutor monitored the messages and whenever a message could be identified as a question containing a key concept, the standardized answer was given; else, the tutor did not join the discussion. Directly following each learning session, a knowledge test was applied to assess participants’ degree of knowledge after the session. The knowledge test consisted of a number of multiple choice items and one open question.

<table>
<thead>
<tr>
<th>Without Turn-Control</th>
<th>Without Referencing</th>
<th>With Referencing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triad</td>
<td>15 (5)</td>
<td>15 (5)</td>
<td>30 (10)</td>
</tr>
<tr>
<td>Quartet</td>
<td>16 (4)</td>
<td>16 (4)</td>
<td>32 (8)</td>
</tr>
<tr>
<td>With Turn-Control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triad</td>
<td>15 (5)</td>
<td>15 (5)</td>
<td>30 (10)</td>
</tr>
<tr>
<td>Quartet</td>
<td>16 (4)</td>
<td>12 (3)</td>
<td>28 (7)</td>
</tr>
<tr>
<td>Total</td>
<td>62 (18)</td>
<td>59 (17)</td>
<td>121 (35)</td>
</tr>
</tbody>
</table>

Table 1: Number of participants (groups) per condition.

Results

The data from the 90 subjects of the prior study and the data of the 31 new subjects were analyzed together (for distribution of subjects to conditions see table 1). The effect of the three factors Referencing, Turn-Control and Group Size was tested with a three factorial covariance analysis with the test score of the knowledge tests as the dependent variable (range 0 to 17) and the test score of the pretest as covariate. The ANCOVA revealed a reliable main effect of Referencing $F(1,111)=9.9, p<.01$. Neither the other main effects (Turn-Control and Group Size) nor the interactions were reliable (see figure 2 for mean and standard deviation of the raw data).

<table>
<thead>
<tr>
<th>Group Size</th>
<th>Without Turn-Control</th>
<th>With Turn-Control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Triad</td>
<td>9.93 (2.48)</td>
<td>10.36 (2.10)</td>
<td>10.09 (2.28)</td>
</tr>
<tr>
<td>Quartet</td>
<td>9.19 (2.32)</td>
<td>10.13 (2.55)</td>
<td>9.66 (2.44)</td>
</tr>
<tr>
<td>Total</td>
<td>9.50 (2.38)</td>
<td>10.23 (2.31)</td>
<td>9.86 (2.36)</td>
</tr>
</tbody>
</table>

Table 2: Mean and standard deviation per condition.

Discussion

The conducted study shows that explicit referencing leads to a higher learning score. Applying a strict turn-control, allowing only one participant to write at a time, has no reliable positive effect.

THE EFFECTS OF DIRECT REFERENCING – POST-HOC ANALYSIS

The post-hoc analysis aims to show evidence for the proposed effects of explicit referencing on the grounding strategies. In the analysis we include only the chat logs of the groups communicating without turn control.

The analysis faces two problems. First, changing the analysis level from individual to group data results in a sample of only 18 discourses, 9 per condition, 5 from triads and 4 from quartets. Because of this small number we cannot use statistical tests. The second problem is the heterogeneity of the discourses. For example, the
The length of the chats varies between 25 and 178 messages for the Without Referencing condition and between 35 and 73 for the With Referencing condition. Therefore, we present only descriptive data, but we think that this illustrates the changes in the communication strategies due to the explicit referencing and gives valuable hints for further research.

Generally, the discourses are topic-centered and task-related. 95% of the learners' messages (947 out of 993) are related to the earthquake-topic, especially to the different aspects of the learning goal. The 5% off-topic messages deal mainly with the coordination of the chat. In 12 out of 18 chats less than two off-topic messages occurred at all.

We propose that explicit referencing changes the grounding strategies, because the explicit relation to the surrounding discourse should reduce the costs for understanding a message (Clark & Brennan). This should affect the participants in two ways:

1. The pressure to keep the messages “near adjacent” should be lowered. That is, participants might be more willing to respond to older messages. As indicators for this, we counted the number of intervening messages between message-response-pairs. The analysis shows that in the discourses without explicit referencing nearly 34% of the messages-response-pairs are adjacent compared to 24% in the discourse with explicit referencing. The number of message-response-pairs with 4 or more intervening messages raises from 18% to 26% (see figure 2).

2. The tendency to send affirmative feedback should be raised. As an indicator for this we counted all messages that had no response, assuming that the affirmative feedback is not again commented. In the discourses without explicit referencing 35% of the messages have no identifiable direct response compared to 43% of the messages in the discourses with explicit referencing (see table 3).

![Figure 2: Mean percentage of messages per message-response-distance](image)

<table>
<thead>
<tr>
<th></th>
<th>Triads</th>
<th>Quartets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without</td>
<td>32.6</td>
<td>37.7</td>
</tr>
<tr>
<td>With</td>
<td>41.0</td>
<td>45.5</td>
</tr>
</tbody>
</table>

Table 3: Mean percentage of messages without response per discourse

Besides the changes in the strategies, we expected that the number of misinterpretations should be lowered. As indicator, we counted all observable situations where a participant misinterpreted a message. The analysis showed that there were quite few such situations: In the discourses without explicit referencing in 4 out of the 9 discourses 9 such situations could be identified, none in the discourses with explicit referencing.

**USAGES OF EXPLICITE REFERENCING – A CASE-STUDY**

In the design of the study presented above, some conditions were quite artificial and do not match the conditions in everyday collaborative situations. The participants of the study had no real interest in the topic, were paid, were assigned randomly to the groups, didn’t know each other, had no or limited common ground, and were
forced to set an explicit reference for each of their contributions. In addition, in one condition they had to follow a defined turn-taking method.

These restrictions to the chat communication are expected to lead to a communication style different from non-restricted communication. For example, we saw in the evaluation of the chat log that some participants circumvented the original idea of the referencing by just setting a reference to the complete material. This is the type of reference that requires the least effort; it is done with a right mouse click at any position in the material. In the following case study, our goal is to find out how groups use the referencing functionality in everyday collaborative situations. We expected that users use references mainly to indicate the respond-to relation between two messages and to a lower extent to point to specific parts in the shared material (deictic use).

Method

We announced a chat session in a team meeting of our research group at Fraunhofer IPSI. The topic of the chat session belonged to ongoing planning work in the research group. The group should generate and discuss scenarios to combine concepts and technology from two projects currently running in our group. These were the DIGITAL MODERATION project, which develops a system to facilitate face-to-face workshops, and the ConcertChat project, which develops chat tools with additional features such as shared material and explicit referencing (for example the tool used in this study). Four research associates volunteered to participate in the chat session with the proposed topic. All participants had an academic background in computer science (3) or information science (1). None had used a chat tool with referencing functionality before.

In comparison to the preceding study, the tool differed in the following aspects: (1) There was no turn-taking regulation. (2) Users were free to use the referencing functionality. (3) They could define multiple references for each contribution, e.g., refer to two or more previous contributions or to refer to contributions and places in the shared material.

Figure 3 shows a screenshot of the tool: In the left part of the window the shared material is presented. The upper right part presents a list of all persons currently online. Below is the scrollable list of all chat contributions and in the lower right part the interface to enter and send a new contribution is located.

Figure 3: The chat tool for the case study. It shows the usage of multiple references for one contribution.
After a short introduction to the usage of the tool to be used for the chat, the group split up into 4 separated rooms each equipped with a networked computer running the chat tool. They filled out a questionnaire about their chat competency and their knowledge about the two projects, which build the basis of the task. Then they read the instructions and background material (1 page) in printed form. Following this, the group started to chat for 50 minutes. The page with the instructions and background material was included in the material area of the chat tool. Finally, the participants filled out a second questionnaire, gathered in one room and reflected on the session together with the authors of this paper.

Results

Chat competence and familiarity with the two projects that form the basis for the task is relatively high: Three participants chat on a daily basis. The fourth participant chats less than once a month. Three participants are involved in the DIGITAL MODERATION project and feel informed about the ConcertChat project. The fourth participant, a new member of the research group, has only minimal knowledge about both projects.

In order to learn more about the usage of the referencing functionality in everyday collaborative situations we analyzed the chat log. For 148 contributions out of the total 193, one or more references have been made to the material section or to other contributions. For 10 contributions more than one reference has been made (8 contributions had 2 references, 1 contribution had 3 references, and 1 contribution had 6 references). In more detail, there were 117 references indicating “response”, 20 deictic references to other contributions, 20 deictic references to the material, and 7 unclear or wrong references.

Contributions with multiple references combine responses, deictic references to other contributions and deictic references to the material in various ways: Combinations for double-reference contributions include 2 responses (1), 2 deictic references to contributions (1), 2 deictic references to the material (1), 1 response + 1 deictic reference to the material (3), 1 response + 1 unclear reference (1), and 1 deictic reference to the material + 1 unclear reference (1). The contribution with three references has 2 responses + 1 deictic reference to a contribution. Finally, the contribution with 6 references has 4 deictic references to contributions + 1 response + 1 unclear reference.

In addition to “responds-to” and the deictic use, we found three ways for using references that we didn’t expect:

- In twelve cases the reference was used to indicate that a previous contribution of the same participant is extended by the current contribution. In one case this connected three contributions, in all other cases it connected two contributions.
- In one case a reference was used as a kind of personal bookmark or pointer to another contribution that contained an explanation but was made 30 contributions before. The contribution contained the text “ah .. I put this here, to find my reference”.
- In one case (the contribution with six references) the references were used to respond and simultaneously collect four ideas generated previously in order to sum up.

The second questionnaire revealed that all participants regard the referencing to other contributions as helpful to improve the chat discourse. Two participants regard the referencing to material as helpful as well. All participants felt stimulated by the possibility to refer to other contributions.

DISCUSSION

The frequent usage of referencing shows that the functionality was accepted by the group. As expected, most of the references (117) indicated a “response” relation, but also 40 references were made to point to other contributions or to the material (20 references each). The participants used multiple relations in 10 cases. And there were unexpected usages of referencing – to connect split turns, to set a bookmark, and to sum up. Only 7 references were unclear or could be recognized as wrong. Altogether, this indicates that the participants quickly learned to use the referencing functionality in a variety of meaningful ways to support the chat discourse.

RELATED WORK

This section consists of two parts: First, we look at tools offering similar functionalities. In the second part we sum up related studies. The main functionalities of the chat tool we used in the studies above were:

- provide chat communication and shared material
- allow single or multiple references to other contributions and/or parts of the shared material.
Tools

From the tool design point of view, there are two principle approaches to ease chat communication: providing references to other contributions (see Threaded Chat and Academic Talk below) and providing references to shared material (see Anchored Conversation and Kukakuka below).

Threaded Chat (Smith et al., 2000) is a chat tool that allows contributions to refer to one other contribution as a “reply-to”. The contributions are presented as a tree, which leads to significant problems, e.g. as new contributions are added to different, potentially distant, branches of the tree. To deal with this problem, our tool presented the contributions in chronological order and represented references as explicit arrows. Another path was taken by Academic Talk (McAlister et al., 2004). It provides two panes in the chat window. One pane presents contributions in chronological order, the other one presents the contributions in logical (tree) order as defined by reply-to relations. In both systems, Threaded Chat and Academic Talk, multiple references and shared material are not supported.

The Anchored Conversations tool (Churchill et al., 2000) allows (normal) chats to be connected to a specific point in a related document. There is no support to refer from one contribution to another or to refer to different parts of the material simultaneously. Kukakuka (Suthers & Xu, 2002) couples threaded discussion and web pages. Multiple references to contributions and references to parts of the material are not possible.

The GraffiDis tool (Leponiemi, 2003) combines relations to texts and graphics. Users enter contributions (which can also consist of graphics and other material) at arbitrary places of the “chat” area. After a certain time the contributions are faded out to the background color. With a “history slider” the user can navigate through the discourse in chronological order. Relations between contributions are indicated by nearby positions in the chat area. References to a contribution are not possible after a certain distance in time as the previous contribution already faded out. Also, multiple references are difficult or even impossible if the contributions to be referred to are at distant positions.

To sum up, none of these approaches support references to contributions, references to shared material, and multiple references.

Studies

To the best of our knowledge, there is no previous study investigating the impact of explicit referencing in chat. But there are two categories of studies that help to interpret our findings. On the one hand, there are studies focusing on how variations of medial features influence the communication behavior and outcomes (McCarthy et al., 1993). McAlister et al. (2004) compare discussions in a normal chat tool with discussions in a tool that forces the participants to choose a sentence opener from a predefined fixed set and allows to explicitly define the reply relation. The discussions in that tool showed more on-topic messages and a higher quality of the argumentations than the discussions in the normal chat tool. Smith et al. (2000) found no differences in the task performance between groups using a normal chat and groups using a threaded chat, but a more balanced participation of fast and slow typists in the threaded chat. Čech et al. (2004) varied different aspects of the chat environment (e.g. size of message editor, availability of chat history) and showed that participants adopt their turn-taking and turn-packaging strategies.

Another relevant category of studies investigate chat usage for solving visual tasks. Suthers et al. (2003) compare face-to-face and chat interactions of dyads building a shared knowledge representation in a shared graphical tool. The study focuses on how the participants manage deictic references to elements of the graph representation. It is shown that the chat groups refer much less to the shared work space and in most cases to recently manipulated items. The authors conclude that online collaborators need better integration of information encountered over time and an easy insertion of visual references to the discussed elements.

Studies of chat use conducted in real world situations show, that experienced chat users discuss in more parallel threads (Isaacs et al., 2002), and that chat users can manage parallel threads when there are “observable contextual relations” (O’Neill & Martin, 2003).

CONCLUSIONS AND OPEN RESEARCH QUESTIONS

Referencing has a positive impact on chat conversations. We expect that this effect is increasing with a growing number of participants in the group as the risk of non-adjacent related turns increases with the number of participants. References are used to express a variety of conversational relations as well as to point to important objects in the chat log or in the shared material. We saw a tendency to focus on one reference per contribution to highlight the most important relation.

As was shown, referencing influences message production as well as message comprehension. This changes also the communication behavior of the participants, especially their strategies for grounding. Referencing as presented here exceeds models of threaded discussions in two dimensions: Messages can have multiple references and referencing can connect a message to other messages as well as to shared material. Thus,
participants can use referencing to express a variety of relation types, not only a reply-to relation as in threaded discussion. By preserving the chronological order and using explicit references participants are still aware of all ongoing threads, which might lead to increased participation in parallel threads. Multiple references might even allow bringing together different threads. Further studies are needed to broaden the empirical basis for the interrelations described above.

In this paper we had only strictly synchronous chat scenarios, i.e. all participants attended the complete chat. It would be interesting to investigate whether referencing has a positive effect also in partly or completely asynchronous scenarios, e.g. supporting latecomers in their comprehension of the missed discussions or learners who use the (referenced) chat log as a learning resource after the discussion has ended.

ACKNOWLEDGMENTS

We thank Ute Linder, Tatjana Koch, and Alexander Schütte for their help in the experimental study, and Torsten Holmer for many fruitful discussions about coherence in chat and for visualizing chat structures.

REFERENCES


