Co-text Loss in Textual Chat Tools

Mariano Gomes Pimentel, Hugo Fuks, Carlos José Pereira de Lucena

Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225
22453-900 Rio de Janeiro, Brazil
{mariano, hugo, lucena}@inf.puc-rio.br

Abstract. The research presented in this article investigates a problem related to the lack of understanding of messages exchanged during chat sessions. There is confusion in the majority of the textual chat tools when various people converse at the same time. Sometimes a participant does not identify the relationship of a new message with a previous one, and is unable to establish a conversation thread—within this research this phenomenon is denominated "co-text loss." The causes, consequences and frequency of this phenomenon are discussed in this article. Two textual chat tools that have been developed to try to reduce the occurrence of co-text loss are presented in this article, together with the results obtained through the use of these tools during synchronous debates among undergraduate and postgraduate students in distance learning courses.

1 Introduction

The research presented in this article began with an analysis of the chat sessions that took place during the Information Technology Applied to Education (ITAE) course [1], which is an on-line course taught by the Computer Science Department of the Catholic University of Rio de Janeiro using the AulaNet learningware [2]. Although enthusiastic about the "different and interesting" activity, the participants of the debates in the ITAE course frequently believe the conversation is confusing: "It is not easy to communicate through such a chaotic tool" (Humberto); "Really chaotic!" (Geraldo); "I liked this debate…however I couldn't follow linearly what was being discussed" (Marcelo) [1].

In a chat with some participants talking at the same time, the result is a tangle of messages where, in many situations, it is difficult to identify who is talking to whom about what—this problem is being denominated here as "co-text loss." and is detailed in Section 2. Two textual chat tools that were developed to try to reduce co-text loss are presented In Section 3. The conclusion is presented in Section 4.

---

[1] In the chat transcript fragments published in this article, the real names of the participants were substituted for pseudonyms. The texts were originally in Portuguese and then translated into English. The original transcripts are available at [3, 4].

2 Co-text Loss

The objective of this section is to define the "co-text loss" phenomenon (Section 2.1) and also to present the investigations about causes (Section 2.2), consequences (Section 2.3) and its frequency (Section 2.4).

2.1 Definition of co-text loss

The initial inspiration for the identification of co-text loss was the perception that the text that resulted from a chat session displayed some features that are similar to text in hypertext: both are non-linear. This similarity spurred an investigation into whether in a chat session a problem similar to the classic hypertext problem also occurred: *disorientation* or *lost in hyperspace* [5]. And, in fact, some participants felt "lost" during chat sessions, as shown in the chat transcript in Text 1.

Text 1. Stating co-text loss. Source: Debate 1 of the ITAE 2000.1 edition (first semester 2000). In this debate a total of 289 messages were produced and sent by 9 participants

```
   24    <Liane> Director, as far as I know it is a piece of
                 authoring software and not Groupware
   26    <Pablo> in my understanding, authoring software contributes
                 to groupware
   30    <Liane> I believe that it is just the contrary, that
                 groupware can help in the authoring process since
                 it can facilitate the communication process
                 between members of a team
 ▶ 31 <Humberto> Contrary to what, Liane, I'm lost
   36    <Liane> When I said the contrary, I didn't mean that
                 authoring defines groupware, but that groupware
                 makes authoring possible
   37 <Humberto> OK
   38    <Pablo> How about both ways, Liane?
```

"Co-text[2] loss" was the term used in this study to designate the phenomenon that occurs in a chat session when a participant does not establish a conversation thread. Co-text loss occurs each time the reader is unable to identify which of the previous messages provides the elements that are necessary to understand the message that is being read. For example, in the transcript fragment in Text 1, it was necessary to identify that message 30 (of Liane) was counter-arguing message 26. Humberto was not able to make this association and expressed co-text loss in message 31.

Co-text loss can be verified through statements such as "what are you talking about?" or "I didn't understand." Such statements are called here *textual manifestations of co-text loss*. It is necessary to emphasize that those statement should

---

[2] "Co-text" designates surrounding text that has been written before or after an enunciation and that provides elements for understanding it. This term is used in Linguistics as an effort to solve the ambiguity of the word context, which has a wider meaning [6].

not always be considered as a manifestation of co-text loss. Upon declaring "I didn't understand," the participant may have identified the co-text of the message but not have understood it for another reason—for example; the participant might be manifesting the inconsistency or non relevance of the argument presented in the message. It is also necessary to emphasize that co-text loss is a cognitive phenomenon—the *textual manifestations of co-text loss* is only one of the consequences of co-text loss, and not the phenomenon itself.

2.2 Causes of co-text loss

From text that is "linear" and "well organized," as generally is the case in books, articles and magazine texts, one expects threading, concatenation, sequence of information. Although a given text may not be merely a chain of enunciations, it is this chain that provides a more legible text. Different than linear and well-organized text, text from a chat session is non-linear. The *linearity* of a chat text is defined here as the percentage of the messages that establish linearity; that is, the percent of the messages associated with the message immediately before them. The *non-linearity* of the chat session is the complement of this percentage.

The predominant non-linearity of a chat session[3] implies features that make it more likely for co-text loss to occur. The greater distance between associated messages[4] makes it more difficult to locate the referenced message, understand the cohesion mechanisms [7] and make the inference of the association between the messages. The non-linearity of chats also make it likely for a confluence of topics to occur: different topics are discussed in parallel, alternately. Although other factors could also be regarded as possible causes of co-text loss, for this research project, it was assumed that the non-linearity of chat sessions is the main cause of co-text loss.

2.3 Consequences of co-text loss

Figure 1 presents a simplified scheme of actions that participants could carry out after detecting co-text loss.

ACTIONS   CONSEQUENCES

CO-TEXT LOSS
1) Rapidly identify the co-text
2) Keep trying to find the co-text ·······► Loss of "conversation rhythm"
3) Give up trying to find the co-text ····► Non-understanding of parts of the conversation
4) Declare co-text loss ···················► Conversation disfluence
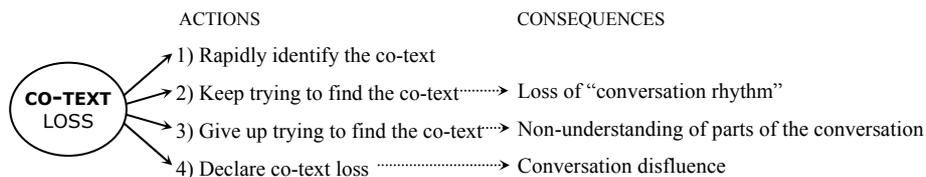
Figure 1. Possible actions after co-text loss detection

Faced with co-text loss, the participant seeks the co-text in the previous messages. If the co-text is quickly identified (action 1 of Figure 1), the conversation continues

---

[3] The debate sessions analyzed on this research project were 85% non-linear [4].
[4] On average, the messages were associated with the 6th previous message [4].

as if nothing had happened. If the participant does not quickly identify the co-text and continues to look for it in the previous messages (action 2 of Figure 1), this search will take up time and effort, and cause the loss of conversation rhythm—while looking for the co-text, the other participants will be continuing the conversation. If the participant stops looking for the co-text, and does not declare its loss (action 3 of Figure 1), she might not understand part of the conversation. If the participant states her co-text loss (action 4 of Figure 1, message 31 of Text 1), another participant may attempt to outline the unidentified co-text (message 36 of Text 1) and, eventually, the participant who lost the co-text may declare her understanding of it (message 37 of Text 1), and the conversation can continue. All of these messages—while necessary for co-text loss repair—cause conversational disfluency [8]: they do not contribute to the development of the topic of the conversation, interrupting the information flow.

The loss of conversational rhythm, the non-understanding of parts of the conversation and the disfluency of the conversation are potential consequences of co-text loss and characterize the phenomenon as a problem.

2.4 Frequency of co-text loss

The frequency of co-text loss situations that occurred during the debates of two ITAE editions was investigated for this research project.
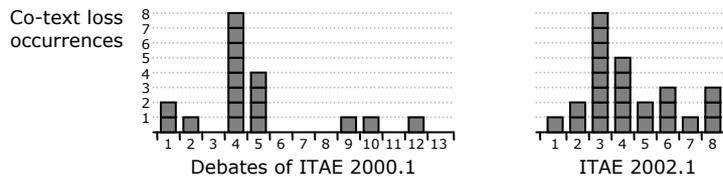

Figure 2. Co-text loss occurrences observed during the ITAE 2000.1 edition. On average 7 participants were present and 336 messages were sent per debate. In the ITAE 2002.1 edition, on average 19 participants were present and 622 messages were sent per debate.

What can be seen, based upon the data presented in Figure 2, is the low average of situations in which co-text loss is declared: 2 situations per debate, 1 situation every 217 messages. However, based upon this data one should not conclude that co-text loss is a sporadic phenomenon. Not all co-text loss is manifested textually.

Another finding of the study was the reduction in the number of situations in which co-text loss was declared. There are two interpretations for this. The former is that over the course of chat sessions the group acquires experience, learns how to better interact and converse. The latter is that the participants learn to better tolerate the phenomenon, reducing their declarations of co-text loss. Even if co-text loss is reduced or becomes more tolerable over the course of the debates, it does not disappear completely, even after a number of chat sessions have been held.

3 Development of chat text tools for diminishing co-text loss

Two chat tools were developed in an attempt to reduce co-text loss and to help understanding the phenomenon: HyperDialog (Section 3.1) and Mediated Chat 2.0 (Section 3.2).

3.1 HyperDialog and the Explicit Threading of Conversation

While the majority of chat text tools organize messages in chronological order, the HyperDialog tool [4] was developed to structure messages by threads. Using this organization, the text sequences are evident. In an isolated thread, the conversation remains totally linear: each message is associated with the message that comes immediately before it. The hypothesis is that the threads mechanism would reduce co-text loss because it imposes structure to the non-linearity nature of the chat, which was identified as one of the main causes of co-text loss (see Section 2.2).

In order to evaluate whether the HyperDialog tool reduces co-text loss, it was put to use in the INformation Technology in Education (INTE) course, a subject taught by the Computer Science Department of the Federal University of Rio de Janeiro. In this course, the HyperDialog tool and a typical chat tool were used in different and intertwined sessions. In this evaluation, 5 debates sessions were held, each one lasting approximately 50 minutes. On average 11 participants were present and sent 173 messages per debate. The debate dynamics of this course was similar to the one of the ITAE course. Contrary to what was expected, there were still instances of co-text loss in the debates in which the HyperDialog tool was used. The data is presented below.
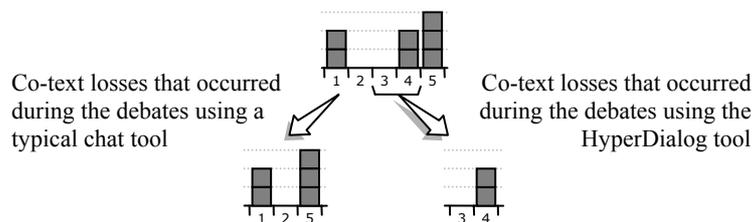

Co-text losses that occurred during the debates using a typical chat tool

Co-text losses that occurred during the debates using the HyperDialog tool

Figure 3. Co-text loss occurrences observed during the INTE 2001.1 edition

Investigating the co-text losses manifested through the use of the HyperDialog tool, it was clear that these occurred to messages that lacked a proper association with the message it referred to. The omission of or error in making this association means supplying the reader with incorrect information—this can disorient the reader regarding the interpretation of the message or generate indisposition towards trying to understand it. When messages are incorrectly associated, the conversation threads become useless.

Analyzing the errors committed by the participants while establishing message associations using the HyperDialog tool, it was seen that approximately 92% of the messages were correctly associated. This percentage indicates that most of time the participants were able to converse while associating their messages during the chat

session[5]. On the other hand, the 8% of the messages that were not properly associated indicated that the participants were having difficulty using the HyperDialog tool—improvements must be made to lower this percentage.

The evaluation of the use of the HyperDialog tool makes it clear that threads help to reduce co-text loss as long as the associations between the messages are correctly established. However, it is also evident that conversation threads make the chats more formal and degrade conversation fluency. This structuring is useful for activities in which understanding of the conversation is highly desirable, as is supposed for debates involving course subject matter. However, the use of threads perhaps is not appropriate for activities involving socialization and recreation, where conversational informality and fluency are highly desirable.

3.2   Mediated Chat 2.0 and the Group Conversation Techniques

The "Mediated Chat 2.0" tool [10] implements the following group conversation techniques: *free contribution*, where any participant can send a message at any time; *circular contribution*, where the participants are organized in a circular queue and, one by one, the first one of the queue can send a message; *single contribution*, where each participant must send a single message at any time; and *mediated contribution*, where only the selected participant can send messages. To take advantage of the conversation techniques, the following debate dynamics was proposed: the moderator was supposed to present the topic (mediated contribution); next, each participant was to send a message commenting on the topic (circular contribution); and then, they were to choose (single contribution) which of the commentary should be discussed in a free way (free contribution). This cycle—topic, comments, vote and free discussion—should be repeated 3 times. The hypothesis is that the use of conversation techniques would reduce co-text loss, as a result of the global organization of the conversation into well defined stages and not by the local organization of the messages that was made possible by the message thread mechanism aforementioned.

In order to evaluate whether the use of the group conversation techniques would reduce co-text loss, the Mediated Chat 2.0 tool was used during the ITAE 2002.2 edition (second semester 2002). In this evaluation, 8 chat sessions were held, with each one lasting approximately 50 minutes. On average, 10 participants were present and sent 364 messages per debate. A typical chat tool was used during the first 4 debates, and the Mediated Chat 2.0 tool was used in the last 4 ones. Co-text lost occurred during this debates according to the data presented in Figure 4.

---

[5] At the beginning of this research project, there was no available data that indicated the feasibility of using threading in chat tools. Currently, it is possible to find some studies about the use of threaded chats [9]. However, these studies show no data regarding the co-text loss phenomenon.
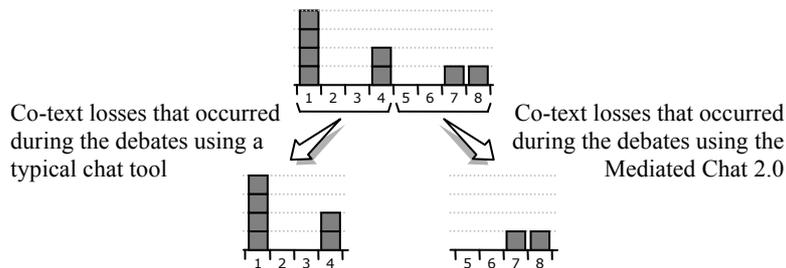
Figure 4. Co-text loss occurrences observed during the ITAE 2002.2 edition.

Investigating the co-text losses that occurred during ITAE 2002.2, it was clear that they occurred only during the free conversation stage. It is precisely during the branching out phase of the free conversation, that other topics started to be discussed in parallel—a trait that makes co-text loss more likely. In the other stages, the conversation thread is easier to be followed.

Although co-text losses still occurred in the ITAE 2002.2 debates, on average only half of the co-text loss situations that had occurred during the debates of the other editions took place. One possibility is that the debate dynamics envisaged to make use of the conversation techniques forces the reduction of co-text loss occurrences. Although the conversation techniques could be enforced without using the Mediated Chat 2.0 tool—by resorting to authority and social protocol—its computational support stimulated the application of the new dynamics to the debate.

4 Conclusion

Textual chat tools have achieved widespread popularity and increasingly people want to use these tools in activities that go beyond socialization and recreation. In this research project the use of chat tools for running synchronous debates in on-line courses was investigated. In these debates, participants frequently complained about confusion during chat conversations. Among the symptoms that provided evidence of such conversational confusion was the recurrence of co-text loss.

The research presented in this article sought to investigate mechanisms that could render better organized chat conversations. In order to investigate these mechanisms, the research sought to look into whether these mechanisms reduced the manifestations of co-text loss. The results that were obtained indicate that the use of threads and the use of conversation techniques help to reduce the occurrence of co-text loss, increasing the understanding of the conversation. However, these mechanisms only partially solve the problem of co-text loss. In future work, other mechanisms such as the use of multiple pages of text to organize the chat conversation, and the imposition of limitations regarding the quantity of messages sent over a given period of time are going to be investigated.

Acknowledgments

References

1. Fuks, H., Gerosa, M.A. & Lucena, C. J. P.: The Development and Application of Distance Learning on the Internet. Open Learning - The Journal of Open and Distance Learning, v. 17, n. 1 (Feb 2002) p 23-38
2. Fuks, H.: Groupware Technologies for Education in AulaNet. Computer Applications in Engineering Education, NY, v 8 issue 3 & 4. Wiley-InterScience (December 2000) 170-177
3. Pimentel, M. G., Fuks, H. & Lucena, C. J. P.: Perda de Co-texto em sistemas de Bate-papo Textual. Technical Report, n. 01/03. Computer Science Department of the of the Catholic University of Rio de Janeiro, Brazil (January 2003)
4. Pimentel, M. G.: HiperDiálogo: ferramenta de bate-papo para diminuir a perda de co-texto. M. Sc. Dissertation. Computer Science Department of the Federal University of Rio de Janeiro, Brazil (April 2002)
5. Conklin, J.: Hypertext: An Introduction and Survey. IEEE Computer (Sept 1987)
6. Crystal, D.: A Dictionary of Linguistics and Phonetics. Blackwell, Oxford (1985)
7. Halliday, M. A. K., Hasan, R.: Cohesion in English. London: Longman (1976)
8. Shriberg, E. E.: Preliminaries to a Theory of Speech Disfluencies. Ph. D. Thesis of the University of California at Berkeley, USA (1994)
9. Smith, M, Cadiz, J. J., Burkhalter, B.: Conversation trees and threaded chats. Technical Report (June 2000) On-line document: ftp://ftp.research.microsoft.com/pub/tr/tr-2000-43.doc [Consultation date: 22nd January 2003]
10. Rezende, J. L.: Aplicando Técnicas de Comunicação para a Facilitação de Debates no Ambiente AulaNet. Forthcoming M. Sc. Dissertation. Computer Science Department of the Catholic University of Rio de Janeiro, Brazil (March 2003)