Source Memorization in Chat Interactions

Gaëlle Molinari, Patrick Jermann, Pierre Dillenbourg, Ecole Polytechnique Fédérale de Lausanne (EPFL), CRAFT, CE 1 630 (Station 1), CH –1015 Lausanne, Switzerland.
{gaelle.molinari, patrick.jermann, pierre.dillenbourg}@epfl.ch

Abstract: This paper reports a study about memorization of online chat interaction. Results show that subjects are very good at recognizing who produced a given utterance, especially if they produced the utterance themselves. Performance was much weaker when subjects recalled who produced the utterance immediately following the given utterance. We investigated several variables in order to predict which utterances are easier to remember.

Introduction

The learning outcomes in collaborative settings are related to the quality of interactions, mostly verbal interactions. Therefore, our CSCL research agenda includes basic questions such as how memory may impact social interaction and collaborative processes. If we take a Vygotskyan perspective, the internalization of conversations requires memorizing them in one way or another. According to Miller and deWinstanley (2002), memory processes are involved in maintaining coherence in conversations. Imagine a person who has difficulty remembering who said what to whom and when during a discussion. This person runs the risk of repeating information and failing to actively collaborate in the co-construction of meaning. The ability to retrieve conversation exchanges from memory is also important in a situation where various people talk at the same time via a chat tool. Pimentel, Fuks and Lucerna (2003) used the terms "co-text loss" to designate the phenomenon that occurs during a chat when participants are unable to establish a conversation thread. According to Horton and Gerrig (2005), the way speakers adapts their utterances to their audience, depends upon the accessibility in memory of the mental representation the speakers constructed about their addressees' knowledge, needs, etc. They also assumed that during interaction, individuals could use their partners as contextual cues to retrieve information they share with them. How do interlocutors memorize their interactions? What characterizes the chat utterances that people remember best? These questions are at the core of collaborative learning mechanisms but haven't received much attention in our community. This contribution reports on investigations we conducted in task-oriented chat discussion between three students.

Method

Studies that investigate conversation memory usually include two phases: a discussion phase and a testing phase (e.g., Keenan, MacWhinney, & Mayhew, 1977). In the first phase, either participants are presented with a written transcription of a discussion, or are asked to participate in a discussion. In the second phase, two types of tests can be used to evaluate whether subjects have memorized either the content of conversation (content memory) or the speaker of a particular utterance (source memory): a recall test or a recognition test. This contribution concerns source memory (our experiment addressed both measures but we focus on the latter). We used a prime-target paradigm to investigate the influence of memory activation upon retrieval. To our knowledge, this paradigm had not been used to investigate interaction memory. During the discussion phase, subjects had to design a working space dedicated to students by taking various constraints into consideration. In our testing phase, participants were first provided with a prime utterance automatically extracted from the transcript of their conversation (a chat). Half of the participants were asked to recognize who said the prime (experimental condition) whereas the other half was assigned to a control condition (no prime speaker recognition). Second, all participants had to recall the speaker of the target utterance, that is, the utterance that immediately followed the prime and then to recall the content of the target. We only report here the results regarding memory for the two speakers (source memory), that is, the prime and the target speakers. The independent variables are the fact that subjects in the experimental group had to recognize the prime speaker as well as several features (see below) used for selecting the primes from the chat conversations.

Hypotheses

Based on Horton and Gerrig studies (2005), we hypothesized that asking participants to identify the speaker of the prime (experimental condition) would facilitate retrieval of the target speaker (H1: *prime speaker effect*).

Two factors frequently identified as influencing memory for conversation are distinctiveness and cognitive effort (e.g., Keenan et al., 1977; Klauer, Wegener, & Ehrenberg, 2002). Memory is higher for conversational sentences that are different in some way (e.g., high interactional sentences). Regarding source memory, it seems that the more distinctive two or more speakers are, the easier it would be to separate their respective contributions. Based on these findings, source memory would be better (a) for conversational utterances that are highly speaker-specific, i.e., for utterances that contain words mostly produced by one speaker during the discussion (H2: *speaker specificity effect*), and (b) for long utterances that usually require more processing effort (H3: *utterance length effect*).

Previous studies on interaction memory also revealed two effects, namely the *order effect* (Igou & Bless, 2003) and the *generation effect* (Miller & deWinstanley, 2002). Primacy (versus recency) effects should occur if participants expect that the most important information should be given at the beginning (versus the end) of the conversation (H4: *order effect*). The generation effect concerns memory for our own contributions during interaction. We thus hypothesized that memory performance would be higher for self-generated than for partner-generated utterances (H5: *generation effect*). This effect – as well as those previously presented (speaker specificity, utterance length, order) – can be explained by the resource allocation hypothesis (Miller & deWinstanley, 2002): Interaction memory essentially depends on the amount and direction of attention collaborators pay to each other.

Apparatus

Thirty male subjects (undergraduate students from the EPFL) participated in groups of three. Each session lasted about one hour, and was composed of two phases: a discussion phase (Phase 1) and a memory phase (Phase 2). In phase 1 (30 min), participants used a text-based chat tool to discuss about the design of a Learning Center that will soon be built on the EPFL campus. Immediately after the discussion, all participants were asked to answer an on-line questionnaire (20 min). They were presented with a series of messages (primes) automatically selected from their chat. In the experimental condition, participants were asked both to read the prime and to identify its speaker, whereas they were only asked to read the prime in the control condition. All participants had then to retrieve both the speaker and the content of the message (target) that was immediately sent after the prime during the chat session.

The 18 primes presented in Phase 2 were automatically selected from the logfiles produced in Phase 1 according to the following rules. All selected utterances were produced in an *intensive context*. The context for a prime is constituted of four sentences preceding it plus four sentences following it. The context intensity is a superficial indicator of whether participants are highly engaged in the chat or not. We defined a high intensity context as containing relatively long messages that are temporally close and produced by all participants. The chat transcript was divided into three parts. Six primes were chosen from each part of the chat. Among six primes in each part, two primes were chosen for each speaker. Finally, among these two primes, one prime was very specific to its speaker, and one was not specific at all to its speaker. Speaker specificity was calculated as the weighted average standard deviations of verbs and nouns usage by the participants. For instance in the sentence “Libraries help”, one noun and one verb are used for the calculation of speaker specificity. Say the verb “help” was used during the chat 10 times by A, 2 times by B and 3 times by C, the standard deviation for this word is 4.36. As the verb was used 15 times overall, the weighted standard deviation is 4.36 / 15 = 0.29. Suppose the noun “library” was used 13 times by A, 12 times by B and 9 times by C, the weighted standard deviation is 2.08 / 34 = 0.06. The speaker specificity for this utterance is the average of weighted standard deviations (0.29 + 0.06) / 2 = 0.175. The rules for choosing prime utterances correspond to 3 variables: the speaker (self versus others), the position (start, middle, end of the chat) as well as the speaker specificity (high versus low). The condition (identification of the prime speaker vs. control) has been also taken into account in our analysis. Due to the many constraints imposed on prime selection, it was often not possible to extract 18 utterances from the chat transcript. In these cases, the memory test was conducted with fewer utterances.

Results and discussion

We examined the effect of our experimental variables on the correctness of both speaker recognition for primes (“who said this?”) and speaker recall for targets (“who said the next utterance?”). Predictions for accuracy were computed through logistic regressions. Because participants worked in groups and responded several times to similar questions, we used mixed effect regressions (responses nested in persons nested in groups) to analyse data.

Concerning the prime speaker recognition, a surprising result is found: accuracy in recognizing the speaker of a given utterance is extremely high (77% correct responses overall; 66% correct when the prime is produced by others and 98% correct when the prime is produced by oneself). Moreover, we clearly observe a *generation effect* (H5): participants always recognize their own messages (β_{self} = 3.71; p = .000). There also seems to be an *order effect* (H4) only for primes produced by others: both primacy and recency effects are observed (utterances produced at the beginning, the middle and the end of the chat: 72% correct, 60% correct and 70% correct, respectively). This effect is however not statistically significant (β_{middle}= -0.42; p = .29). We believe that such an effect might appear with a longer delay between the conversation phase and the memory test. Finally, speaker recognition performances tend to be higher for low speaker-specific primes. Although this result is not significant ($\beta_{specificity}$= -1.5; p = .21), it contradicts our hypothesis H2 that assumes a positive effect of *speaker specificity* on memory for chat interactions. Possibly our definition of speaker specificity could be enhanced to include signs and expressions (e.g., sms style abbreviations) that were not recognized by the part-of-speech tagger used in automatic analysis of the logfiles.

Concerning target speaker recall, the subject's average performance is weaker but still above the threshold of 1/3 that corresponds to responses by chance since we have three speakers (42% correct responses overall). The *generation effect* (H5) is again observed (50% correct for self-produced targets and 38% correct for targets produced by others); this difference is however marginally significant (β_{self} = 0.47; p = .07). We identified a detrimental effect of the *prime length* (H3) on target speaker recall ($\beta_{wordlength}$ = -0.058; p = .018): it seems to be easier to recall the speaker of a target when it is preceded by a short rather than a long message in the chat session. This effect suggests a resource allocation problem. Longer is the prime, more cognitive effort is needed to process the prime and less participants in conversation pay attention to the subsequent message. Neither the chat position nor the speaker specificity does significantly affect recall accuracy. Our results also do not support the hypothesis that people use their partners in conversation as retrieval cues (H1). Indeed, asking participants about the prime speaker does not have any effect on target speaker recall. The low number of participants could explain why all these effects did not emerge. Moreover, as Pimentel et al. (2003) said, there is not always a relation between a message and the one that immediately precedes it in a chat session. That could be another reason of the absence of prime effects.

Conclusion

This study deals with a highly important aspect for CSCL, that is, the memorization of chat utterances. The main result is the fact that group members are very good in recognizing who said what. Recall performance is however much weaker when subjects have to remember the speaker of utterances that follow a given prime. This highlights the difficulty for collaborators in remembering and reacting to what their partners state in the chat discussion. Performance is especially weak for target messages preceded by long contributions. We interpret this result as resulting from a resource allocation problem. Our results do not confirm the assumption that partner-specific information could be used as retrieval cues to improve memory because it might be more accessible in memory (Horton & Gerrig, 2005). The results underline the necessity for chat systems to provide users with tools that help them to differentiate their partners' contributions. In future research related to memory for chat interactions, the effect of other variables – for instance, the duration of a chat session or the delay between chat and testing – should be taken into account.

References

Horton, W. S., & Gerrig, R. J. (2005). Conversational common ground and memory processes in language production. *Discourse Processes*, *40*(1), 1-35.

Igou, E. R., & Bless, H. (2003). Inferring the importance of arguments: Order effects and conversational rules. *Journal of Experimental Social Psychology*, *39*, 91-99.

Keenan, J. M., MacWhinney, B., & Mayhew, D. (1977). Pragmatics in memory: A study in natural conversation. *Journal of Verbal Learning and Verbal Behavior*, *16*, 549-560.

Klauer, K. C., Wegener, I., & Ehrenberg, K. (2002). Perceiving minority members as individuals: The effects of relative group size in social categorization. *European Journal of Social Psychology*, *32*, 223-245.

Miller, J. B., & deWinstanley, P. (2002). The role of interpersonal competence in memory for conversation. *Personality and Social Psychology Bulletin*, *28*(1),78-89.

Pimentel, M. G., Fuks, H., Lucerna, C. J. P. (2003). Co-text loss in textual chat tools. *4th International and Interdisciplinary Conference on Modeling and Using Context, LNAI 2680.* Stanford, CA, USA, June, pp. 483-490.