Proximity and View Awareness to Reduce Referential Ambiguity in a shared 3D Virtual Environment

David Ott, Pierre Dillenbourg

Geneva Interaction Lab - TECFA
University of Geneva (Switzerland)
David.Ott@tecfa.unige.ch , Pierre.Dillenbourg@tecfa.unige.ch

ABSTRACT

This contribution investigates the role of virtual space on social interactions during collaborative tasks. We previously observed that MUD users rely on spatial positions to refine the conversational context and thereby facilitate mutual understanding. Supporting mutual understanding is a main challenge of CSCL research. We explore how this may happen in a continuous 3D space (VRML). Our first hypothesis was that the proximity of the emitter to the referred object clarifies the referential context. Our second hypothesis stated that the receiver uses gaze awareness in order to guess which object the emitter refers to. The experiment results confirmed the first hypothesis, surprisingly rejected the second hypothesis and reveal complex interactions between the two.

Keywords
Virtual space, virtual reality, collaboration, proxemics, view awareness

INTRODUCTION

Many CSCL environments are based on a spatial metaphor: a virtual campus includes rooms, buildings, etc. Why ? Besides the motivational role and the navigational role of the spatial metaphore we think that virtual space would play a functional role in collaboration among peers. This hypothesis arose from findings of a series of studies using a text-based virtual reality (a MOO environment) to create the task that subjects had to perform (Dillenbourg & Traum, 1999). We observed that the rate of acknowledgment was significantly higher when two subjects were located in the same virtual room (Dillenbourg, Mendelsohn & Jermann, 1999). Moreover, the delay of acknowledgment was significantly shorter in virtual co-presence. The implicit assumption behind these observations is that subjects pay attention to spatial awareness information. In another experiment (Montandon, 1996) using a task that requires spatial coordination, we suppressed the spatial awareness messages that are automatically generated by a MOO environment. The subjects compensate this information loss by performing specific user locating commands. These results reveal a mechanism that goes beyond co-presence effects: if there is a clear structural mapping between the problem space and the virtual space, reasoning on mutual positions is reasoning on collaboration strategies, or, in other words, *spatial awareness supports coordination at the task level*. We observed another functional role of space. The room that also includes one agent or key object seemed to be used by subjects as the by-default context to disambiguate utterances. This observation may develop our functional understanding of virtual space: we hypothesize that *spatial awareness supports grounding by providing subjects with the contextual cues necessary to refer to objects*. Understanding how a virtual environment may facilitate the construction of a shared understanding is a key challenge of CSCL research and the main goal of the experiments reported here. These preliminary observations were bound to the room paradigm of MUD environments. Would our preliminary observations be confirmed in a more systematic experiment using a continuous space, i.e. in space where rooms do not simply define in/out relations, but where distance matters?

METHODOLOGY

We constructed an experimental 3D VRML Virutal Environment (VE) where two subjects (N=20 pairs) are required to collaborate to solve a simple object-matching task. The subjects are seated in different computer rooms and can only interact with their partner through the VE. The multi-user 3D VE constructed for use in the experiment is figurative, and poor in details. The task (10 randomized rounds) is for both subjects to locate a target object from amongst nine objects located in the VE, to communicate their (the emitter's) finding to the partner using a structured communication interface and then for the partner (the receiver) to confirm or reject the proposition. During the task the target object is always shown in the upper portion of the viewpoint. All of the nine objects are cuboids, and are highly similar to each other; therefore the object-target matching task is far from straightforward. A quick glance at objects in the VE is insufficient to ascertain a match with the target object, subjects must, rather, take time to explore the objects in detail. The representations of the subjects in the virtual space, i.e. their avatar, are simple red cones. While a user explores the VE his avatar moves accordingly in the VE. Each user sees the avatar of his/her partner (or can decide to look at it), but being inside their own avatars the subjects cannot see themselves. The use of simple upright cones, as avatars, was a crucial experimental choice, as this representation carries no information on the orientation of the avatar. Therefore, there is no way for a user to tell the field of view of the partner on the VE. We provided the users with a view awareness tool: every object in the field of view

of the partner's avatar is highlighted using a different color to those objects out of their field of view. The presence or absence of this awareness tool constitutes the experimental condition of the study. Position and orientation of the avatars in the VE are logged every second. Avatar actions, such as the manipulation of objects, or communication using the structured interface, are also logged. From this raw data we computed several measurements (dependent variables) like various distance measures of the emitter to the reference object and an ambiguity measure consisting of the sum of examined objects by the receiver prior to his answer, i.e. the greater the number of manipulated objects the greater the ambiguity of the situation. We postulate the following hypotheses:

- *Hypothesis 1: The proximity of the emitter to the referred object clarifies the referential context.* We think that although there is no explicit way for the emitter to reference the target object, the emitter still can use a collaborative feature of space, i.e. proximity, to identify the reference object. The nearer the emitter to the referenced object the less ambiguous is the referential context for the receiver.

- *Hypothesis 2: The 'view awareness' clarifies the referential context.* By providing the view awareness tool we think to facilitate the receiver's task. Sequences with view awareness should be less ambiguous and therefore have a clearer referential context.

- *Hypothesis 3: The distance from the emitter to the referred object should increase with the 'view awareness'.* According to (Clark and Wilkes-Gibbs, 1986, cit.in Clark & Brennan, 1991) 'least collaborative effort' principle, conversing partners tend to minimize their collaborative effort. The redundancy of context disambiguating clues, i.e. proximity and view awarenss, should lead to a slackening of the collaborative effort when possible, that is, the proximity to the object. In conditions with view awareness the emitter will tend to be more distant from the referenced object.

RESULTS & CONCLUSION

Results show that distance is positively correlated to the number of examined objects. Though the correlations are relatively small, three out of the four distance measures are highly significant. Thus, we consider hypothesis 1 to be confirmed. We didn't observe any difference between sequences with or without view awareness (p=.983). Hence, the second hypothesis is invalidated. Finally, although, distance measures to the referred object tend to be greater in the condition with the awareness tool, an ANOVA revealed no significant interaction between view awareness and proximity.

In conclusion, we say that users may use some features of virtual space, namely distance, to support a core mechanism in collaboration, defining the referential context. It still remains an open issue for us to dissociate to which extend the emitter's move to the object was due to the task constraints or reflect a deliberate deictic move. It only indicates that, when the emitter *has to* perform this move for task-specific constraints, then the receiver is able to use this information to disambiguate references. This information may however be used by CSCL designers for instance to decide how they position objects in virtual space.

BIBLIOGRAPHY

Dillenbourg, P., Mendelsohn, P. and Jermann, P. (1999). Why spatial metaphors are relevant to virtual campuses? In J. Levonen & J. Enkenberg (Eds.) Learning and instruction in multiple contexts and settings. Bulletin of the Faculty of Education, 73. University of Joensuu, Finland, Faculty of Education.

Dillenbourg, P. and Traum, D. (1999). Does a shared screen make a shared solution ? International Conference on Computer Supported Collaborative Learning, Stanford, décembre 1999.

Montandon, L. (1996) Etude des mécanismes de coordination spatiale dans un environnement virtuel de collaboration. Mémoire de Diplôme d'Etudes Supérieures en Sciences et Technologies de l'Apprentissage. Non publié. TECFA, Faculté de Psychologie et des Sciences de l'Education, Université de Genève.

Clark, H. H. & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. Levine & S. D. Teasley (Eds), Perspectives on Socially Shared Cognition. APA, 1991.