Towards a computational model of frame of reference alignment in Swedish dialogue

Simon Dobnik*, Christine Howes* and John D. Kelleher†

*CLASP
University of Gothenburg, Sweden
simon.dobnik@gu.se
christine.howes@gu.se

†School of Computing
Dublin Institute of Technology, Ireland
john.d.kelleher@dit.ie

Abstract

In this paper we examine how people negotiate, interpret and repair the frame of reference (FoR) in online text based dialogues discussing spatial scenes in Swedish. We describe work-in-progress in which participants are given different perspectives of the same scene and asked to locate several objects that are only shown on one of their pictures. This task requires participants to coordinate on FoR in order to identify the missing objects. This study has implications for situated dialogue systems.

1. Introduction

Directional spatial descriptions such as "to the left of green cup" or "in front of the blue one" require the specification of a frame of reference (FoR) in which the spatial regions "left" and "front" are projected, for example "from where I stand" or "from Katie's point of view".

A good grasp of spatial language is crucial for interactive embodied situated agents or robots which will engage in conversations involving such descriptions. These agents have to build representations of their perceptual environment and connect their interpretations to shared meanings in the common ground (Clark, 1996) through interaction with their human dialogue partners.

There are two main challenges surrounding the computational modelling of FoR. Firstly, there are several ways in which the viewpoint may be assigned – *intrinsic* (assigned by the reference object of the description (Levinson, 2003)); *extrinsic* (an external viewpoint such as superimposed grid structure); or *relative* (with reference to a conversational participant or object in the scene).

The second challenge is that the viewpoint may not be overtly specified and must be recovered from the linguistic or perceptual context. Such underspecification may lead to situations where conversational partners fail to accommodate the same FoR leading to miscommunication.

There are a number of factors that affect the choice of FoR, including: task (Tversky, 1991), personal style (Levelt, 1982), arrangement of the scene and the position of the agent (Taylor and Tversky, 1996; Kelleher and Costello, 2009; Li et al., 2011), the presence of a social partner (Duran et al., 2011), the communicative role and knowledge of information (Schober, 1995), but very little work has investigated choice of FoR from a dialogic paradigm, in which participants can work through potential misunderstandings together using processes of repair.

We are interested in how participants align their spatial representations in dyadic text dialogues when they perceive a scene from different perspectives. How do they identify if a misalignment has occurred, and what strategies do they use to get back on track?

2. Method

Task Using 3D modelling software we designed a virtual scene depicting a table with several mugs of different colours and shapes placed on it. As shown in Figure 1, the scene includes three people on different sides of the table. The people standing at the opposite side of the table were the avatars of the participants (the man = P1 and the woman = P2), and a third person at the side of the table was described to the participants as an observer "Katie".

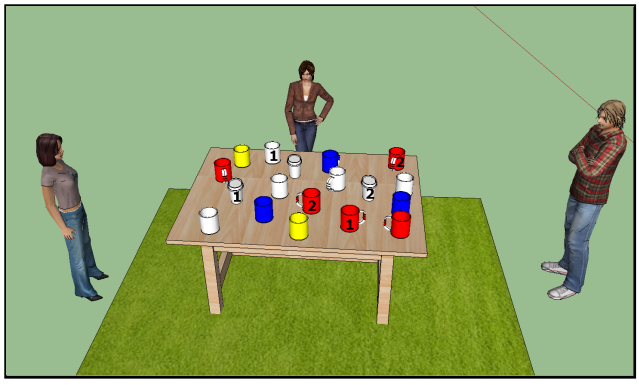

Figure 1: A virtual scene with two dialogue partners and an observer Katie. Objects labelled with a participant ID were removed in that person's view of the scene.

Each participant was shown the scene from their avatar's point of view (see Figures 2 and 3), and informed that some of the objects on the table were missing from their picture, but visible to their partner. Their joint task was to discover the missing objects. The objects that were hidden from each participants are marked with their ID in Figure 1.

Procedure Each participant was seated at their own computer and separated so that they could not see each other or each other's screens. Communication was through an online text based chat tool (Dialogue Experimental Toolkit, DiET, (Healey et al., 2003)), which records each key press and associated timing data.

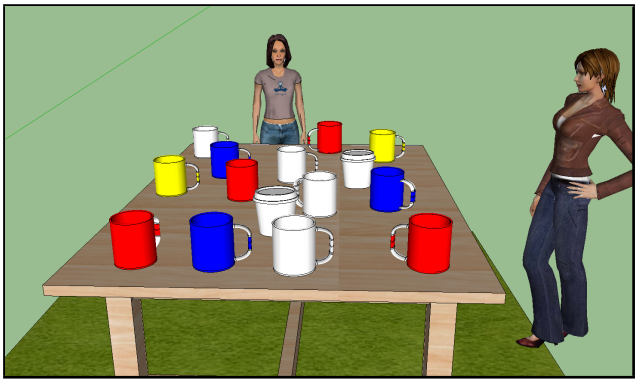

Figure 2: The table scene as seen by Participant 1.

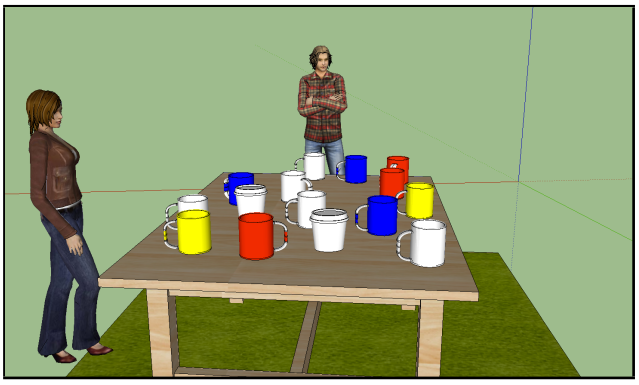

Figure 3: The table scene as seen by Participant 2.

3. Summary of results

A pilot study (Dobnik et al., 2015) suggests that there is no general preference of FoR in dialogue but the choice is related to the communicative acts of particular dialogue games at specific points in the dialogue. There is also evidence that participants align their FoR locally over a sequence of turns, but not globally; at points of misunderstanding it may be prudent to shift FoR in order to get the conversation back on track. We isolate several conversational games where the dynamics of the FoR assignment appears to be linked to other properties of interaction between the agents, for example whether they are focusing on a particular part of the scene or whether they are identifying individual objects scattered over the entire scene. It follows that alignment is consistently used as a strategy but there are other factors that trigger changes in FoR.

In this work-in-progress we look at Swedish dialogues, to investigate if these findings hold cross-linguistically, when resources for resolving misunderstandings may not be the same across languages. We also examine whether a selection/change of the FoR could be predicted from the (textual) dialogue data. We hypothesise that dialogue turns contain sufficient information about the dialogue games that conversational participants are engaged in and to which the FoR assignment appears to be linked.

Through quantitative data analysis we attempt to identify features that are predictive of FoR changes and which would be useful for annotating and extending our corpus described above. The overall goal is to provide a training dataset for machine learning that would allow us to build a model of FoR assignment. Finally, we also investigate the suitability of different machine learning models for the task.

Acknowledgements

Thanks to Kim Demaret for collecting the Swedish data.

References

Herbert H. Clark. 1996. *Using language*. Cambridge University Press, Cambridge.

Simon Dobnik, Christine Howes, and John D Kelleher. 2015. Changing perspective: Local alignment of reference frames in dialogue. In Christine Howes and Staffan Larsson, editors, *Proceedings of SemDial 2015 (goDIAL): The 19th Workshop on the Semantics and Pragmatics of Dialogue*, pages 24–32.

Nicholas D. Duran, Rick Dale, and Roger J. Kreuz. 2011. Listeners invest in an assumed other's perspective despite cognitive cost. *Cognition*, 121(1):22–40.

Patrick G. T. Healey, Matthew Purver, James King, Jonathan Ginzburg, and Greg J. Mills. 2003. Experimenting with clarification in dialogue. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, Boston, MA, Aug.

John D. Kelleher and Fintan J. Costello. 2009. Applying computational models of spatial prepositions to visually situated dialog. *Computational Linguistics*, 35(2):271–306.

Willem J. M. Levelt. 1982. Cognitive styles in the use of spatial direction terms. In R. J. Jarvella and W. Klein, editors, *Speech, place, and action*, pages 251–268. John Wiley and Sons Ltd., Chichester, United Kingdom.

Stephen C. Levinson. 2003. *Space in language and cognition: explorations in cognitive diversity*. Cambridge University Press, Cambridge.

Xiaoou Li, Laura A. Carlson, Weimin Mou, Mark R. Williams, and Jared E. Miller. 2011. Describing spatial locations from perception and memory: The influence of intrinsic axes on reference object selection. *Journal of Memory and Language*, 65(2):222–236.

Michael F. Schober. 1995. Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about locations? *Discourse Processes*, 20(2):219–247.

Holly A. Taylor and Barbara Tversky. 1996. Perspective in spatial descriptions. *Journal of Memory and Language*, 35(3):371 – 391.

Barbara Tversky. 1991. Spatial mental models. *The psychology of learning and motivation: Advances in research and theory*, 27:109–145.