Adaptability of Communicative Resources in AR-based Cooperation

Christian Schnier [1], Karola Pitsch [1], Angelika Dierker [1,2], Thomas Hermann[2]

[1] Applied Informatics Group & [2]Ambient Intelligence Group
Faculty of Technology, CITEC, Bielefeld University, Germany
cschnier@techfak.uni-bielefeld.de

Abstract

We present an initial investigation from a semi-experimental setting, in which an Augmented Reality (AR) system, based on Head Mounted Displays (HMD), has been used for real-time collaboration in a task-oriented scenario (design of a museum exhibition). While allowing for a range of technical augmentations, the setting also restricts – due to the wear of HMDs – the participants' 'natural' communicational resources. Our analysis reveals that – under these particular conditions – some everyday strategies of establishing co-orientation with the co-participant turn out to be not functional for the participants. At the same time, we find that some participants change their referencing strategies to overcome system-based limitations and to develop a – under these particular conditions – more efficient method in orienting the co-participant to specific objects or to the interaction situation itself: Participants transform their individual deictic gestures on several objects into other forms of gestural activities like for example the lifting or the tilting of an object. These particular changes in object trajectory are done in order to orientate the co-participants and establish joint attention. Furthermore, gestural referencing seems to be highly variable and contextual, if important interactional resources are artificially reduced.

Index Terms: gesture, multimodality, Augmented Reality, co-orientation

1. Introduction

Over the last 15 years a range of initiatives has emerged that develop and explore Augmented Reality (AR) systems, in which the user's perception of the world is overlayed with additional, digital information [1, 2]. Most commonly, these systems focus on augmenting the user's visual perception by video taping in real-time the user's environment and displaying this image together with overlayed additional information on a screen. To achieve this effect, existing AR-systems either (i) exploit the cameras/displays of recent mobile phone technologies or (ii) equip the user with specialized glasses, so-called head-mounted displays (HMD). The first approach benefits from using an already available technology and easy integration into the user's everyday practices, which is reflected in the current boom of applications for navigation, interactive tourist guides etc. The second approach allows to support richer and more complex activities, during which the users could freely use their hands to manipulate objects, which is relevant e.g. in aircraft maintainance where 3D construction plans are made available, in situ, to the engineer. Whilst existing research predominantly focuses on individual users, little is known about AR-technologies in collaborative settings.

Dierker et al. [3] have developed a HMD-based Augmented Reality setup for collaborative scenarios as an interaction research tool which we here apply to investigate the participant's methods of orientation with reduced communicative resources while allowing for controlled interactional conditions. In comparison to natural face-to-face interaction, to wear HMDs and to see the world through its lenses results in limited access to usually available communicational resources: reduced field of view [4], lower resolution [3] and problems in determining the co-participant's focus of visual attention [5]. Additionally, looking through HMDs results in significantly less eye rotation and increased head orientation when attempting to focus on a given point or object [6]. These changed conditions of interaction raise some empirical questions with regard to methods of orientation: How can co-participants, under these conditions, organize their interaction and coordinate their activities? How can they establish joint attention?

In this paper, we will present some findings from a quasi-naturalistic AR-experiment, in which we have equipped pairs of users with HMDs and asked them to jointly design a museum exhibition while arranging a set of objects (the exhibits) on a given floor plan. Our analysis will address the questions raised above and – using sequential micro-analysis stemming from Conversation Analysis – focus (1) on the practical problems in orientation (2) the emergence of interactive methods for orientation.

2. Background

While a few collaborative AR-setups have been proposed [7, 8, 3], little is known yet as to how participants can deal with the conditions implied by the technical constraints when attempting to fulfill a joint task. In fact, with the collaborative HMD-based AR-scenario, a new prototype of face-to-face communication seems to arise: On the one hand it encompasses aspects typical of face-to-face interaction: physical co-presence, shared interaction space, participants can touch, smell and hear each other and they can jointly manipulate the same objects. On the other hand, participants see the world through the eyes of a video-camera with a reduced field of view and resolution and – due to cost and computing power – mostly monoscopic vision, and virtual augmentations are not necessarily similar for both co-participants (cf. [9]); these features are comparable to technologically mediated settings. Empirical investigation of such systems has revealed the limitations of such technologies in comparison to unmediated face-to-face interaction: Yamashita et al. state that "gaze, gesture and other body movements are generally not as effective as in normal face-to-face communication" [10]. And Luff et al. notice: "the system fractures the environments of action and inadvertently undermines the participants' ability to produce, interpret, and coordinate their actions in collaboration with each other" [11].

On the level of system development, these limitations could be seen as a set of technical challenges to be addressed. From an interactional perspective, we can begin to use this system – in combination with additional components (see below, system description (3)) – as a sophisticated research tool that enables us to investigate a range of interactional

phenomena more closely. In such a co-present, but technologically mediated setting, the co-participants are faced with the task of organizing their multimodal interaction [12]: to coordinate their actions [13], to monitor and take into account the co-participant's current state of action [14] and to establish joint attention [15].

Moreover, it seems to be an interesting approach in our particular setting to investigate the range of different signaling systems in adopting – under certain conditions – different interactional functions. Goodwin (2003) shows with his concept of "ecology of sign systems" that in a patient with aphasia disturbed signaling systems can be substituted. These signaling systems can, if necessary, fulfil different functions. Thereby the functional adaptation is expressed by the term "ecology". "The term ecology is used to note the way in which these separate systems function as differentiated, interdependent components of a larger whole that can adapt to changing circumstances" [16]. While the process of adaptation is difficult to investigate in aphasic deseases because of long lasting communicative adjustments (long-term adaptation), it immediately becomes online observable within our controlled setting (short-term adaptation).

Therefore, with our setup it is also much easier to investigate and analyze the fundamental factors and aspects that are responsible for different adaptation strategies of the participants within the course of their interaction.

3. AR-System for collaborative task-oriented interaction

Over the last years, Dierker et al. [3] have developed an AR-system that allows for real-time collaboration of two users and to record and intercept the users' natural communication channels. The system encompasses the following components (cf. Fig. 1): HMDs with an integrated camera that captures the view from the user's perspective and passes it on as a video frame that is projected on the screens of the corresponding HMD. Similarly, audio signals can be captured with microphones and relayed via in-ear headphones. This paradigm allows to precisely record the relevant sensory information available to interacting users. This enables us to reconstruct the users' audio-visual perceptions and to gain a better understanding of their respective member's perspective in co-present interaction. We furthermore record the detailed head movements by inertial sensors worn on the head. This allows us to measure accurately amplitude, frequency and timing of head gestures such as nodding and head shaking.

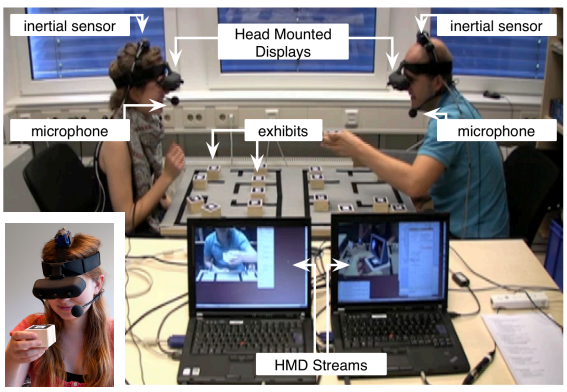


Fig. 1: System for collaborative AR-use

Secondly, we augment virtual objects on top of physical objects in the interaction space using ARToolkit [17]. Specifically, for the museum planning scenario, these virtual objects are exhibit pictures shown on wedge-shaped 3D objects. As novel contribution we introduced a coupling of users by a joint-attention-support channel: each user sees by coloring of the exhibit objects whether and how much they are in the view field of their interaction partner. More precisely, the object's frame color changes from yellow (peripheral) to red (in the center of the partner's field of view) (cf. [3]).

4. Experiment: Collaborative museum exhibition design

To investigate collaboration under the specific conditions of AR-technology, an experiment has been conducted (08/2010) in which pairs of users were asked to jointly plan a museum exhibition while arranging a set of objects (the exhibits) on a given floor plan [18, 19, 20, 9]. The participants were seated face-to-face at a table, equipped with AR-glasses, microphone headsets and an inertial sensor on top of their heads. The participants were asked to carry out three subsequent tasks: (1) In a familiarization phase (5 minutes) the participants were asked to chat with their partner about a self-chosen topic in order to familiarize with being videotaped and wearing the devices. (2) In the following individual phase (with vision obstructing barrier) they were asked to individually plan a museum exhibition using a set of 8 different exhibits each (wooden blocks as material 'handles' for augmented objects sitting on top of the blocks), and arrange them on a given floor plan. (3) In the subsequent dyadic phase, the participants were asked to discuss their arrangement of the exhibits with their partner (without barrier) and to develop a joint solution for all 16 exhibits in one of the two equi-oriented identical floor plans.

5. Co-orientation as a practical problem: Establishing coordination and joint attention

The specific conditions of our AR system induce practical problems in establishing coordination and joint attention for the participants. We have presented a detailed analysis of these specific conditions as part of a system evaluation [9]. Our analyses revealed three particular conditions, which are substantially different from daily interactional conditions known from face-to-face communication:

1. *Dual ecology and the world's instability:* Our AR-setting presents a dual ecology for the participants: On the one hand, they have physical access to the real world such as the table with wooden blocks and the floor plan, and are in physical co-presence with the interaction partner. On the other hand, on the level of visual perception, this world is mediated through the display of the real-time video-stream and the added virtual augmentations. During the interaction, the AR environment is indicated by a mixture of stability and instability. While the wooden blocks and the plan determine stable reference points, this is not valid for augmented objects.
2. *(Dis-)Embodiment:* The dual ecology between real and virtual world and the lack of a stereoscopic view also influence the ways in which participants can deal with their own and the co-participant's bodily existence in the world. This becomes evident in co-coordinated activities like grasping, e.g. when the participants exchange objects with their co-participant.
3. *Orientation and interactional coordination:* The participants have – due to the AR-glasses – only a highly reduced field of view with a masked periphery, so that

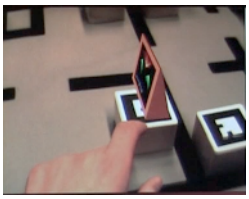
they can either look at their co-participant or inspect (parts of) the museum plan and/or exhibits. Thus, focusing on the task, they are hardly aware of the partner's physical representation, body movements, head orientation etc.

Let's consider a first fragment F1 from one of the dyads from our experiment, which occurs early in the third interaction phase (32s after its beginning). It gives a first insight into the practical problems that the participants face when attempting to establish mutual orientation on a given object on the museum plan and demonstrates point 1 and 3 of the above mentioned particular interactional conditions.

Fragment 1, (VP08, 32:53-33:00)

```
01 B: |äh=also diese LASERshow hier;|(0.2)
       euh=well this LASERshow here;

02 A: |hm:=hm:-|
   B: |und die|(0.2)|siehste hier,|(0.5)
       and the        you see it here,
                                  #1a+b
```



#1a (B´s field of view)

#1b (A´s field of view)

```
03 B: |in diesem RAUM hab ich
       den extra reingestellt,|
       I put it in this ROOM for a good reason,

04 B: |<<p>mh=weil>|
       mh=because
   A: |warte,      |
       please wait,
```

At the beginning of this fragment, participant B tries to orient his co-participant A to the object 'lasershow', which he attempts to do exclusively by verbal means: "`euh-well this LASERshow here;`" (01). Additionally he fixates the object, which he is talking about. Because of the highly reduced field of view his co-participant A is not aware of his head orientation. Referencing by the focus of attention is thus not possible (cf. point 3). However, the verbal deictic reference turns out to not be sufficient for a precise reorientation of the co-participant. Participant A does not re-orient to look at the object, but answers with an elongated "`hm:=hm`". Participant B treats this as only 'claiming', but not 'showing' understanding. This can be seen because B correctly treats her answer as not having followed his orientation and reformulates his suggestion. He adds "`you see it here,`" (02). This time, he adds a gestural pointing to the object (#1) and designs his turn as a question, which projects the co-participant's confirmation. Comparison of participant A's versus participant B's field of view reveals that the object in participant A's field of view is actually not augmented (cf. #1a vs. #1b). However, the object determines a stable reference point in participant B's field of view (cf. #1a). Other cases demonstrate also a different alignment of the objects on the visual markers of the wooden blocks or indicate marker flipping, i.e. participants might perceive different augmentation on exactly the same object (cf. point 1). However, after B's reformulation, participant A reacts to this second attempt by leaning forward, thereby receiving the correct augmentation, and commenting on the indicated object. Thus, we should ask about how participants establish a common orientation under these complicated conditions at all. This will be analyzed in the following section.

6. Emerging interactive procedures for establishing co-orientation

The following fragments shed some light on how orientation processes of participants change in the sequential course of interaction. In the run-up of the presented sequences the generation of attention to specific objects with deictic pointing occurs frequently. Deictic gestures in combination with deictic expressions e.g. with object deixis and local deixis, are commonly used as a procedure to incite a co-participant to orient to a particular place/object/etc. in face-to-face communication. This orientation process, however, turned out not to be a sufficient tool in the modified interaction condition, as presented here, to orientate the co-participant (cf. fragment 1). This is because with the use of HMDs the peripheral view is limited or even completely missing and deictic activities in the visual periphery are often not perceivable.

6.1. Establishing new interactional procedures

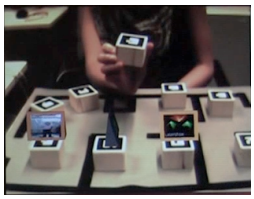
Fragment 2 shows how participants change their gestural activity and interactively develop particular methods of orientation, which they appear to consider as appropriate in the contextual circumstances. We contrast this fragment with fragment 3 from another group of participants in which the same method of orientation was tested as in fragment 2. However, this method was not consequently continued because in this case - due to technical problems - this method was not effective and is explicitly considered as a problem of orientation by the co-participants.

Fragment 2, (VP08, 35:41-35:53)

```
01 A: |jetzt äh hab ich hier noch so=eine|
       now euh I have here also such a

02 A: |plasmascheibe ne,|
       plasma dial right,
   B:                    |okay;|
                          #1a+b
```

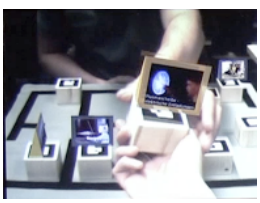
#1a (A´s field of view)



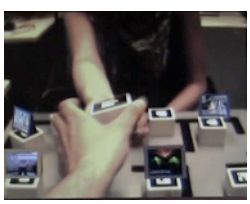
#1b (B´s field of view)

```
03 A: |die WIRKT natürlich auch besser (-)|
       which APPEARS of course much better

04 A: |im dunkel (-)|
       in the dark (-)
   B:                |darf ich die mal sehen,|
                      may I take a look at it,
                                      #2a+b
```



#2a (A´s field of view)



#2b (B´s field of view)

At the beginning of the fragment participant A refers to the object "`plasmascheibe`" (plasma dial) by verbal means

(01+02) and highlights it by lifting it in the air (cf. #1a+b). The attached tag-question "ne," (right,) plays the role of a "securing of notification" or in this particular case a kind of reassurance in order to indicate a successful orientation. Participant B reacts on this via an "okay;" and, thus, indicates that he realizes the object to talk about. Actually, the object "Plasmascheibe" (plasma dial) is not augmented in B's field of view during this point in time (cf. #1b). We observed that the participants immediately indicated orientation problems at the beginning of the interaction and tried to eliminate these problems via particular repair-mechanisms. However, in the course of the interaction the indication of orientation problems is often delayed. Participants adapt their communicative behavior to the technical problems of the system and wait, if the augmented object perhaps is presented a little delayed.

Also in this fragment the "non-augmentation" is not considered as an orientation problem by participant B. Thus, participant B does not explicitly indicate the "non-detection" of the object, but solves the orientation problem by asking for the object (04, picture #2a+b). After participant B has grasped the object, he could manipulate it by changing the angle of inclination. In this, the marker becomes detectable to the system and the virtual object is correctly displayed on the cube.

6.2. Approved procedures become a stable routine

In the further course of interaction the above-mentioned orientation process (cf. 6.1) is more frequently observable and becomes a stable orientation routine. Figure 2 shows the timeline of orientation methods in the collaborative dyadic phase of "VP08". During the first three minutes deictic pointing is the dominant orientation method.

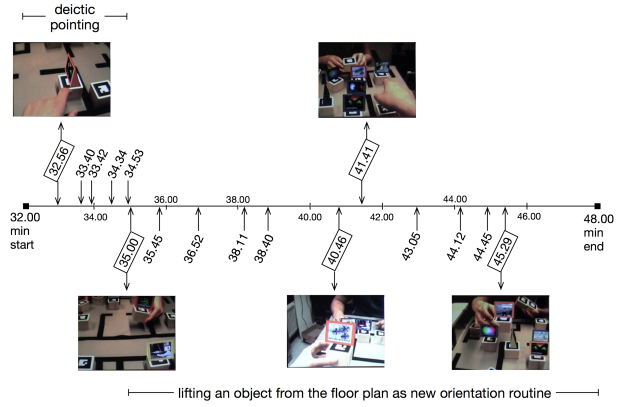


Fig. 2: Timeline of orientation methods

As shown in Fragment 1, these orientation method turns out not to be sufficient for a precise re-orientation of the co-participant. Instead, subject A initiates a new procedure in 35:00. She lifts the object and presents it slightly elevated for a longer stretch of time. We observe that this orientation-method levels off in the further course of interaction and becomes a stable routine for orientation, which turns out to be much more sufficient than pure deictic pointing procedures.

7. Applying the orientation method does not necessarily become a routine

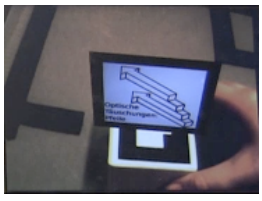
Fragment 3a also shows – equivalent to fragment 2 – that a participant lifts an object for a better orientation and presents it to the co-participant. Although the method of orientation is identical to the procedure used by VP08, for VP09 there is no stabilization as a routine. The following fragment illustrates that the way in which the co-participants react to a initiated new procedure influences the continued existence of the procedure.

The beginning of Fragment 3a shows that the object "optische Täuschung" (optical illusion) is explicitly offered as next relevant object (01). In this early phase of the fragment both participants are focused on different parts of their own map (cf. #1a+b). Shortly after participant A lifts the object and marks her action as a new orientation method simultaneously by verbal means: „ich heb das mal grad hoch, (-)" (I just lift it up) (03) participant B uses this announcement for reorientation and changes her focus of attention and looks to the raised object (cf. #2a+b).

Fragment 3a, (VP09, 31:01-31:14)

01 A: |ähm hatt ich hier noch anzubieten::|
ah I have to offer here

02 A: |optische täu=täuschungen steht hier bei|
optical illusions stands here near
#1a+b

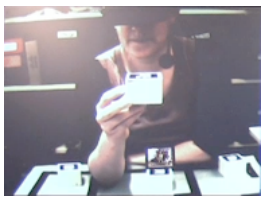


#1a (A´s field of view)

#1b (B´s field of view)

03 A: |mir=ich heb das mal grad hoch,(-)|
with me=I just lift it up
B: |äh:::|
er:::
#2a+b

#2a (A´s field of view)



#2b (B´s field of view)

04 B: |ja dann kann ich=es aber nicht sehen;|
well but then I can´t see it;

05 B: |also (.) dann kann die=die kamera|
I mean then the=the camera
A: |aso|
ok
#3a+b

#3a (A´s field of view)

#3b (B´s field of view)

06 B: |!JA! GENAU;| |ahja okay;|
!YES! RIGHT; okay;
A: |das hier;|
that one;

Picture #2b shows that the object is not indicated for participant B. Because of this participant B reacts with an extended hesitation signal "äh:::" (er:::) and then gives an indication in 04 that A's way of orientation leads to the result that the object is not displayed on her display. Participant A reacts upon this by slightly tilting the object in direction of the co-participant and by bringing it closer to the co-participant. Even if this method leads to a sufficient orientation of her co-participant, it can be seen that from now on the "optical illusions" are not augmented in her field of view (cf. #3a+b).

Now participant B clearly indicates in 06 that the orientation is successful by the phrases "!JA! GENAU;" (!YES! RIGHT;) and "ahja okay;". Noticeable in this group of participants is that this particular way of orientation is not further used in the course of interaction. Obviously and in comparison to fragment 2, the explicit treatment of the orientation problem seems to lead to an avoidance of this strategy in the further course of interaction.

The orientation via lifting an object is not further tracked by this group of participants, which happens to be the case because of problems in orientation. However, it emerges that precisely the element, which finally lead to a successful orientation in the "problem case" of fragment 3a, is more frequently used for orientation in the course of interaction. Thereby, orientation is achieved via changing of the angle of inclination in direction of the co-participant. In this case, fragment 3b shows a way of orientation which incorporates the familiarization of fragment 3a and which is adopted as a fundamental routine of orientation by the use of this efficient element.

Nevertheless, the lifting of the object does not completely stop in the periods of time after fragment 3a. However, we could observe that the lifting does not act as an orientation function like in fragment 2. This becomes manifest when analyzing the run-up of fragment 3b.

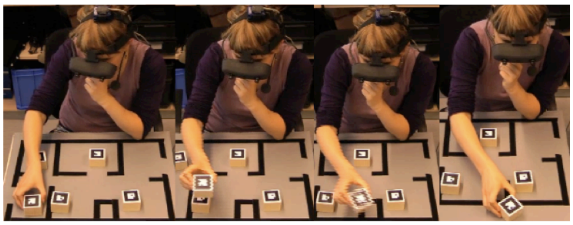
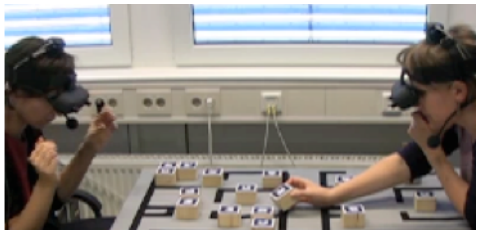


Fig. 3: Run-up of fragment 3b (32:28-32:31)

In the run-up of this fragment participant A considers whether the object „Löschen einer Kerze durch Paukenschlag" ("dousing a candle by a drumbeat") could be placed onto a particular object of participant B's plan. Thereby, participant A lifts the object in the direction of participant B's plan and places it between both plans. The trajectory of participant A's hand is without any gradient and, thus, indicates a clear difference to the observed trajectory used in fragment 2 for the co-participant's orientation. Additionally, we should realize that the object-manipulation is not accompanied by a verbal output. The missing deictic expressions, in particular, indicate that participant A does not try to orientate participant B to the lifted object. In fact, participant A manipulates the object for an individual preparation phase in order to be able to operate from a new position (cf. Fig.3). The hypothesis of an individual preparation instead of a cooperative acting is corroborated by the fact that participant B does not perceive her action. It can be shown in fragment 3b, that the intention to orientate the co-participants to the specific object takes place later in time.

Fragment 3b, (VP09, 32:48-32:56)

```
01 A: |die SEIFENbla:sen und ähm und hier der|
       the BUBBLES and ah and this
                                  #1
```



#1 slightly object tilting for orientation

```
02 A: |die paukenschlag(.)kerze könnten noch|
       drumbeat-candle could be

03 A: |zusammen ne,|
       put together right,
   B:                |oh=JA das: äh:: scheint|
                      oh=YES that er:: seems
         #2#3

04 B: |mir auch;|
       reasonable to me;
```

#2 (B´s field of view)

#3 (B´s field of view)

Participant A orients the co-participant to the object with the particular deictic expression „hier" (here). Additionally, she changes the angle of inclination (01, #1) and audible touches upon the table top with the wooden block. This gives a further signal for orientation. "Touching upon the table top" might be the attempt to ensure the demanded orientation. Picture #2+#3 shows that the former slight change of the angle of inclination results in a further and stronger forward-backward-motion of the wooden block. This procedure indicates that the instability of the augmented world becomes understandable and that the participants try to bypass the complicated interactional conditions by modifying their gestural activity. In accordance with fragment 2 an orientation-reinsurance is made by an attached "ne," (right,).

7.1. Conclusion of the results

Comparison of both fragments 2 and 3 reveals that the group of participants develop different methods to successfully orient the corresponding co-participant. With regard to the concept of "ecology of sign systems" by Goodwin, we could observe that the interacting participants adjust their signalling systems according to the contextual conditions during the course of interaction. Thereby, a successful method for orientation, i.e. a method that is accepted by the interacting participants, depends on the particular way of handling the corresponding method. In the beginning both example cases reveal the same method for orientation and in both example cases this method leads to orientation problems, i.e. the augmented object is not displayed. While this lacking orientation is not an explicit problem in fragment 2, some

explicit formulations emerge that indicate the corresponding method of orientation as "insufficient" in fragment 3. In this latter case the method is further adjusted until it forms a stable orientation function. Thereby, this adjustment takes place on the basis of obtained experiences, which have established in the unfamiliar environment. It emerges that the used element of lasting orientation (slighly tilting the object in direction of the co-participant and by bringing it closer to the co-participant) is already part of the orientation method of "Lifting an Object".

8. Discussion

In this paper we have shown that gestural activities of participants with a high functionality in orientation adapted to the contextual conditions. Common interactional patterns from face-to-face communication like for example deictic pointing on objects transform to other forms of gestural expression. Nevertheless, these altered forms have the same function, i.e. they serve as orientation mechanisms for the co-participants. Thereby, the process of adaptation is based on interactive negotiation. Basically, this process of interactive negotiation takes place by verbal means, which is frequently observed in our particular setting. Whether gestural activities become a stable routine for orientation or whether it is necessary to make further adjustments - as shown in fragment 3a/b - depend on the participants' way to cope with the orientation problems. While in fragment 2 an existing orientation problem is solved on implicit request for a particular object, fragment 3 shows an explicit verbalization of the problem. Current gestural orientation strategies are not used any longer and are not going to be modified in the further course of interaction if there is any problem of orientation. Different types of gestures emerge to take the same function with regard to gestural activity. In this view, an accurate classification of gestures in terms of their functionality might be expected not to be very effective because of the sensitivity of gestural activities to particular contextual circumstances.

Furthermore, our study indicates that the open question of what determines a particular gestural activity might be solved with the help of a multimodal view from conversation analysis. The lifting of an object in our setting is functionally equivalent to a deictic gesture and is processed by the recipient as a kind of intention for orientation. In this view, gestural activities are those bodily actions, which evoke an interactive effect, i.e. they exhibit a fundamental role in actions of communication.

9. Acknowledgements

This work has been kindly supported by the Collaborative Research Centre 673 Alignment in Communication, funded by the German Research Foundation (DFG).

10. References

[1] Caudell, T. and Mizell, D., "Augmented Reality: An Application of Head-Up Display Technology to Manual Manufacturing Processes", Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences, vol. 2:659-669, 1992.

[2] Azuma, R., "A Survey of Augmented Reality", Presence: Teleoperators and Virtual Environments, 6(4):355-385, 1997.

[3] Dierker, A., Bovermann, T., Hanheide, M., Hermann, T. and Sagerer, G., "A multimodal augmented reality system for alignment research", Proceedings of the 13th International Conference on Human-Computer Interaction, pp. 422-426, 2009.

[4] Arthur, K., "Effects of field of view on performance with head mounted displays", Dissertation, University of North Carolina, 2000.

[5] Brennan, S., Chena, X., Dickinsona, C., Neidera, M. and Zelinsky, G., "Coordinating cognition: The costs and benefits of shared gaze during collaborative search", Cognition, 106(3):1465-1477, 2008.

[6] Kollenberg, T. et al., "Visual Search in the (Un)Real World: How Head Mounted Displays Affect Eye Movements, Head Movements and Target Detection", ACM SIGGRAPH conference proceedings, 2008.

[7] Schmalstieg, D., Fuhrmann, A., Szalavari, Z. and Gervautz, M., "Studierstube – An Environment for Collaboration in Augmented Reality", CVE ´96 Workshop Proceedings, 1996.

[8] Billinghurst, M. and Kato, H., "Collaborative Augmented Reality", Communication of the ACM, 45(7):64-70, 2002.

[9] Schnier, C., Pitsch, K., Dierker, A. and Hermann, T., "Collaboration in Augmented Reality: How to establish coordination and joint attention?", European Conference on Computer-Supported Cooperative Work, 2011.

[10] Yamashita, J., Kuzuoka, H., Yamazaki, K., Miki, H., Yamazaki, A., Kato, H. and Suzuky, H., "Agora: Supporting Multi-participant Telecollaboration, Proceedings of HCI, pp. 543-547, 1999.

[11] Luff, P. et al., "Fractured ecologies", Human-Computer Interaction, 18(1):51-84, 2003.

[12] Goodwin, C., "Action and embodiment within situated human interaction", Journal of Pragmatics, 32(10):1489-1522, 2000.

[13] Deppermann, A. and Schmitt, R., "Koordination. Zur Begrüßung eines neuen Forschungsgegenstandes", in R. Schmitt [Ed], Koordination, Gunter Narr Verlag Tübingen, 2006.

[14] Goodwin, Marjorie H., "Processes of mutual monitoring implicated in the production of description sequences", Sociological Inquiry, 50(3):303-317, 1980.

[15] Kaplan, F. and Hafner V., "The challenges of joint attention", Interaction Studies, 7(2):135-169, 2006.

[16] Goodwin, C., "The Semiotic Body in its Environment", in J. Coupland and R. Gwyn [Ed], Discourses of the Body, pp. 19-42, New York, 2003.

[17] Kato, H. and Billinghurst, M., "Marker tracking and hmd calibration for a video-based augmented reality conferencing system", Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, pp. 85-94, 1999.

[18] Pitsch, K. and Krafft, U., "Von der emergenten Erfindung zu konventionalisiert darstellbarem Wissen. Zur Herstellung visueller Vorstellungen bei Museums-Designern", in U. Dausendschön-Gay, Ch. Domke and S. Ohlhus [Ed], Wissen in (Inter-)Aktion, pp. 189-222, de Gruyter, Berlin, 2010.

[19] Luff, P., Heath, C. and Pitsch, K., "Indefinite precision: the use of artefacts-in-interaction in design-work", in C. Jewitt, The Routledge Handbook of Multimodal Analysis, pp. 213-224, 2009.

[20] Dierker, A., Pitsch, K. and Hermann, T., "An augmented-reality-based scenario for the collaborative construction of an interactive museum", Bielefeld: Bielefeld University, 2011.