

# Negotiating Spatial Goals with a Wheelchair

**Thora Tenbrink**

I5-[DiaSpace]

SFB/TR 8 Spatial Cognition  
University of Bremen

tenbrink@sfbtr8.uni-  
bremen.de

**Shi Hui**

I5-[DiaSpace]

SFB/TR 8 Spatial Cognition  
University of Bremen

shi@informatik.uni-  
bremen.de

## Abstract

We present our iterative approach to enabling natural dialogic interaction between human users and a wheelchair, based on the alternation of empirical studies and dialogue modelling. Our approach incorporates empirically identified conceptual problem areas and a dialogue model designed to manage the available information and to ask clarification questions. In a Wizard-of-Oz experiment employing the first version of the model, we test how verbal robotic reactions can enable users to provide the information needed by the wheelchair to carry out the spatial task. Results show that the output must be extraordinarily coherent, temporally well-placed, and aligned with the user's descriptions, as even slightly deviating reactions systematically lead to confusion. The dialogue model is improved accordingly.

## 1 Introduction

Most advanced work on dialogue systems focuses on human-computer interaction scenarios in which either the user requires information from an expert system (e.g., Kruijff-Korbayová et al. 2002), or the user and the system negotiate a joint task such as making reservations (Rieser & Moore 2005), or the system engages in tutoring the user within a specific area of interest (Clark et al. 2005). In such tasks, there are typically no particular complications with respect to time or space: Although the dialogue takes place in real time, there are no fundamental context-related effects of temporal

delay or spatial mismatch. Complementing this research, there is a growing interest in dialogue systems employed in real time in spatially embedded interaction scenarios, such as situated human-robot dialogue. Such scenarios typically employ robots designed to accomplish service tasks for users instructing them by using natural language. Work in this area often focuses on a number of specific techniques designed to overcome the particular complexity of such a situation (e.g., Lemon et al. 2003, Spexard et al. 2006, Kruijff et al. 2007). Our own work fits into this latter endeavour by focusing on the spatiotemporal matching problems that are typical for a dynamic setting. Our users are involved in the process of reaching a spatial goal together with the robot in a wayfinding setting. The particular challenge in our framework lies in reaching mutual agreement in relation to the actual surroundings in spite of the fact that humans' and robots' spatiotemporal concepts differ in crucial respects.

Related work also focusing on route descriptions is addressed, for example, by the Instruction-Based Learning group (e.g. Bugmann et al. 2004), and by MacMahon et al. (2006). Our current focus is on a detailed qualitative analysis of the discourse flow between human and robot, using a realistic interaction scenario with uninformed users that is tailored to the actual technological requirements. This particular approach is not to our knowledge adopted elsewhere (though see Gieselmann & Waibel 2005 for a different scenario), but is specifically needed to establish and improve the relationship between implemented functionalities and humans' intuitive reactions at being confronted with an autonomous transportation device. In this paper, we first describe our approach including earlier empirical re-

sults and a sketch of the first version of our dialogue model. Then we present the results of another empirical study testing the model, discuss the ensuing improvements, and conclude by outlining the next steps in this iterative process.

## 2 Previous work

One of the prominent aims in the SFB/TR 8 Spatial Cognition (Bremen/Freiburg)<sup>1</sup> is to enable smooth and efficient spatiotemporally embedded language-based interaction between humans and robots. For this purpose we explore uninformed users' natural preferences in tasks resembling the future functionalities of our robots in basic respects, coupling technological development with empirical investigations. In the long run, our system will implement ontological knowledge as described in Hois et al. (2007), the development of which is also based on our targeted empirical findings, in addition to a careful examination of the existing literature on spatial language semantics and usage (Tenbrink 2007). Our dialogue system architecture is described in Ross et al. (2005). While the system itself is not restricted to application in a particular robot, we focus here on an application with the autonomous wheelchair "Rolland" (Lankenau & Röfer 2000). In Shi & Tenbrink (forthc.), we describe the first steps in adapting the system for an indoor route description scenario. The main focus in that work is on matching the users' spontaneous utterances with the robot's implemented *conceptual route graph* (Krieg-Brückner & Shi 2006). In the following we summarize the results.

### 2.1 Empirical results

Our first empirical study was designed to collect spontaneous utterances and examine users' generalized strategies in a scenario resembling the targeted robotic task. Our users were told to move with the robotic wheelchair through an office environment and describe a range of places and locations to the robot. After that, they were asked to instruct the robot to move to one of these places.

From the collected natural language data, we extracted the following potentially problematic consequences of our users' linguistic choices and

strategies. Most typically, the utterances may contain a reference to an object or location in the real world that the robot is incapable of resolving. This may be due to the vocabulary available to the robot, to the name tags attached to objects and locations in the robot's internal map, to the user's employment of a different expression than that expected by the robot, or to the robot's inability to establish the exact spatial relationship that the user refers to. The latter point is enhanced by the fact that natural spatial utterances are typically underspecified (Tversky & Lee 1998); they only point to a vague spatial direction that needs to be matched to other knowledge sources, and they often lack information about a required ingredient (such as the *relatum*). Since the robot's perceptual abilities differ from the human's, there is a high potential for mismatches especially in the (normal) case of underspecification. On top of that, the utterances we collected in our scenario reflect a high degree of uncertainty on the part of the users.

A different problem is that users are unsure about the level of granularity or detail suitable for the instruction. Some instructions directly refer to the goal location, while others only give directional information such as "straight on – to the left". Since the robot has access to higher-level information, this method is not efficient as it leads to a continuous need for interaction. Also, to match instructions with the implemented conceptual route graph, the robot needs information about spatial boundaries of the route segments, which is often not provided, at least not explicitly. The information provided by the users is also often too vague to be matched to the robot's knowledge.

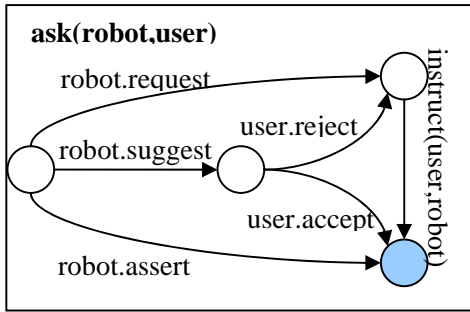
### 2.2 Dialogue system

The first version of our proposed dialogue model was designed to deal with each of the identified problem areas. In the case of reference resolution problems, underspecification, and missing boundaries, the robot asks for more information. If a conflict between the description and the robot's internal map is detected, the robot makes an assertion to inform the user about this disparity. In case of ambiguities, the robot may provide a suggestion to the user. These ideas were integrated within a dialogue model based on the CONversational Roles model (COR) (Sitter & Stein 1992). Figure 1 shows a depiction of a clarification subdialogue *ask(robot, user)*, initiated by the robot, a part of the dialogue

---

<sup>1</sup> Funding by the DFG is acknowledged. Also, special thanks to Kerstin Fischer who was crucially involved in the preparation of the empirical work reported here.

model. In the diagram *request*, *reject*, *accept*, *suggest* and *assert* are dialogue acts, while *instruct* (*user,robot*) is another subdialogue within the dialogue model. Circles represent dialogue states; the marked one is the final state. The subdialogue *instruct*(*user,robot*) may involve iterative processes such as those described by Clark & Wilkes-Gibbs (1986), in which the agreement on a particular kind of reference may take several turns.



**Figure 1** Clarification subdialogue

Our examination of the collected data shows that our formal model should theoretically capture the majority of the potential communication problems identified on the basis of the (monological) first study. In order to account for dynamic dialogue processes, and to put the dialogue model to the test, we conducted a second study in which the robot reacted verbally to the users' utterances. This is particularly important since our cases of clarification relate neither directly to the semantic nor the pragmatic level of understanding (cf. Schlangen 2004), but rather, to the cognitive domain: the robot needs to know precisely how the users' cognitive representation should be matched to its own internal conceptual map. Therefore, standard clarification mechanisms such as various forms of *reprises* or *clausal*, *constituent*, or *lexical* clarifications (Purver et al. 2003) do not readily apply in this particular situationally embedded domain of interaction. Our second study is presented next.

### 3 Empirical investigation

17 German and 11 English native speakers participated in this experiment. The setting in this second study exactly matched that of the first, except that in this case, in the second (route instruction) phase a human "wizard" was seated behind a screen who triggered prefabricated robotic utterances following a certain schema based on the dialogue model

developed before. The schema was devised based on our knowledge about the range of variability in the users' spontaneous utterances, as gained from the first study. The wizard's instructions were as follows: If the user simply states the goal location or reference to a room without providing further information, the robot informs the user that this location is unknown to it, and *requests* further information. If the user provides an underspecified spatial direction such as "then left", the robot *suggests* a precise location to turn according to its internal knowledge, or *requests* clarification in a number of predefined ways, formulated so as to induce the speaker to provide the relevant information on a suitable level of granularity. These reactions account for those cases in which boundaries cannot be inferred from probable interpretations of combined utterances (which should often be possible at least to a certain degree). The wizard could also assume a representation mismatch and react by letting the robot *assert*: "Sorry, this does not match with my internal map". Thus, using a range of preformulated utterances, the wizard was able to produce a reasonably natural dialogue with the user without natural language generation while sounding "automatic" as suitable for the robot. The design was intended to presume a high amount of mismatch and need for clarification (Fischer 2003).

As before, the linguistic data were recorded, transcribed, segmented into TCUs, ("turn constructional units", cf. Selting 2000),<sup>2</sup> and analyzed using the methodology of a detailed qualitative discourse analysis. In particular, since we are interested in the cognitive elements and spatial information content, we categorized the route instruction data with respect to whether each TCU contained:

1. a **directional** or motion-based term, such as "straight on" or "turn left"
2. a reference to a spatial **location**: either a landmark (or sub-goal) or the goal itself, e.g., "go to the office"
3. a reference to a **path entity** ("the hallway").

These distinctions were further examined with respect to whether the landmarks or (sub-)goals in 2. as well as the path entities in 3. were spatially anchored as in "the office *on the left*" or "the first

<sup>2</sup> TCUs are defined on the basis of interactionally relevant completion, taking syntactic, semantic, pragmatic linguistic evidence as well as activity-related factors into account.

hallway *on the left*", and whether they occurred together with a path-describing term such as "*past the office*" or "*down the hallway*". These aspects reflect insights on basic elements of route descriptions (e.g., Denis et al. 1999, Gryl et al. 2002). A specific spatial segment could be described in full by combining all three categories: "go straight on down the hallway in front of you towards the third office on your right". However, most TCUs contain only parts of this information. Other parts may be expressed in or inferable from adjacent TCUs. The component analysis serves here for a first evaluation of the data, though they cannot capture the intricate diversity of the users' distinctions (cf. Klippel et al. in press). More detailed annotations are possible and desirable for our subsequent work, for instance, integrating qualitative and quantitative distance and orientation information (which plays a minor role for the present analysis). In addition to the component-related analysis (section 3.1), we pursued a procedural approach by analysing the development of particular stretches of discourse in detail. In sections 3.2 through 3.4 we present the generalized results of this analysis in relation to the utterance categories presented in 3.1 along with a number of examples.

### 3.1 Component analysis

In the 11 English and 17 German dialogues, we identified 123 English and 244 German TCUs produced by the user and directed to the robot. In the following, we do not differentiate further between the languages since we do not focus on language differences here. 103 (less than one third) of the TCUs were non-spatial (such as answering "Yes" to a robot's question).

Of the remaining 264 TCUs, about one third (31.44%) were purely direction-based (category 1). Almost half of the 264 TCUs (45.08%) contained mention of a location (category 2); one-third of these were neither spatially anchored nor part of a path-description; rather, they presupposed knowledge of the mentioned place, as in "go to the stugaroom". Almost one-third (29.55%) of the 264 TCUs contained a path entity (category 3), half of these occurring together with a path-describing term, and one-third specifying the path entity spatially. Thus, as with landmark entities, the identity of path entities is typically presupposed. We now turn to a closer look at a number of samples in order to investigate the success of our first dialogue

model reflected by our wizard's reactions. Here we did not establish abstract categories for measuring dialogic success, as we felt this would not do justice to the flexibility of the discourse processes.

### 3.2 Smooth dialogue with clarifications

Some of the dialogues turned out to be entirely unproblematic, they appeared to be completely covered by the dialogue model and did not exhibit any communication problems. Here is an example (D1) (numbers in brackets refer to pauses in seconds; the examples segment the utterances here into idealized turns rather than TCUs).

ROBOT I'm ready.  
 USER please go to the stugaraum  
 ROBOT Oh, I'm sorry, I must have missed that location. Can you please explain to me how to get there?  
 USER leave this room (1) then (2) turn left (2) follow the hallway (2) and then take (1) another left  
 ROBOT Is this the first possibility to turn?  
 USER yes (2) then continue left continue down that hallway  
 ROBOT Up to which point do I go straight?  
 USER you go straight until the last room on the left side  
 ROBOT Okay, I think I got it. I am taking you to the Stugaroom now.  
 USER okay go  
 ROBOT I am sorry, my batteries are too low. Please recharge them first.

This example shows how the speaker first assumes full knowledge on the part of the robot, by simply referring to the goal location (category 2). The clarification question induces the speaker to provide more specific information structured in segments according to relevant decision points, employing categories 1 and 3. Then the robot suggests that the left turn is the first possibility to turn, which is confirmed. Utterances of category 2 (here: "continue down that hallway") are underspecified with respect to the end boundary, motivating another clarification question by the robot. The user gives the information, and the robot is satisfied.

### 3.3 Shift of granularity level

Some utterances do not reflect a level of granularity that is suitable for the robot. Our preformulated utterances in some cases successfully motivated a shift in this respect as intended by the form of the robot's requests, as shown in the following example D2.

USER to go to the stugaroom (1) turn (4) 180 degrees (2) exit (1) room three four four zero. turn (1) left (3) go straight (1) ahead

ROBOT Up to which point do I go straight?  
 USER to the - I don't know (5) go straight (1) turn left to the (3) to the copy room (3)  
 ROBOT When I'm left of the room with the photocopying machine, what do I do?  
 USER you go straight until (1) you get to the stugaroom

The user starts out with category 1 descriptions accompanied by hesitations and pauses, omitting information about the boundaries of actions. After the robot's first question the user starts to mention places (category 2) but seems to be uncertain about the correct spatial relations. The robot's second question then seems to encourage the user so that the last part of the route instruction is provided fluently and on a level of granularity more directly matching the robot's internal spatial knowledge (combing categories 1 and 2). – In the following example D3, the robot asserts confusion with respect to its internal map:

ROBOT tut mir leid, aber das passt jetzt gar nicht zu meiner internen Karte. (I'm sorry, but this does not match with my internal map).  
 USER hm  
 ROBOT bis wohin fahre ich geradeaus? (Up to which point do I go straight?)  
 USER bis zur zweiten Linkskurve (until the second left turn)

Here, the robot's request for a spatial goal serves to remedy the confusion and is readily taken up by the user.

### 3.4 Spatiotemporal congruity

As exemplified by the previous examples, our pre-formulated answers turned out to be successful whenever they fit sufficiently to the user's descriptions by precisely filling in gaps or underspecifications. However, these clarifications need to be precisely temporally adjusted; as Rieser & Moore (2005:245) point out, "clarification should not be postponed, and immediate, local management of uncertainty is critical." Our example D4 illustrates how a very slight delay may cause confusion and non-clarification of the underspecified item.

USER you turn left, go straight, turn...  
 ROBOT Up to which point do I go straight?  
 USER after the copyroom continue straight

The question about when exactly to cease going straight remains unanswered; the user essentially blends the next route segment (which involves going straight) with the robot's question. This confusion is due to the users' choice of rapidly sequencing category 1 utterances that leave boundaries underspecified, which according to the

dialogue model triggers the robotic reaction of explicitly asking for boundaries.

Getting back on track is extremely difficult once the discourse flow has been interrupted in an unsuitable way. This may lead to confusions in the represented spatiotemporal sequence, as in the following example D5.

USER wenn wir aus dem Raum A 3440 heraus fahren biegen wir links ab fahren dann geradeaus (coming out of the room A 3440 we turn left and then go straight)  
 ROBOT wo soll ich abbiegen? (where should I turn?)  
 USER links. dann an der nächsten Abbiegemöglichkeit nach links biegen wir dort ab (left. then at the next possibility to turn left we turn).

Here, the robot's question is probably intended to refer to the user's description of *going straight*. But the user mentally goes back to the previous expression *turn left*, and then returns immediately to the point where he was interrupted. Matching this kind of non-sequential information to the robot's internal map is certain to cause severe problems. In this case, the user's interpretation of the robot's clarification question could probably have been avoided if the robot had acknowledged the user's description so far (by saying "Okay" prior to asking the question), so that the user knows the question refers to the current route segment rather than a previous one.

### 3.5 Discussion

Our analysis of utterance components shows that a substantial amount (one third) of speakers' spontaneous route descriptions towards the robot were based purely on spatial directions, rather than providing information about the boundaries of a route segment or the location of a spatial goal or sub-goal (landmark). Taken by itself, this result is similar to our monologic study reported in Shi & Tenbrink (forthc.) where the proportion of purely directional TCUs is nearly 40%. Such instructions are informative when given together with additional information in adjacent turns (Tversky & Lee 1998). However, the robot may not always be able to integrate this information suitably, given the implemented features of the conceptual Route Graph. Also, some of our participants relied entirely on underspecified directional information, leading to the need to infer implicit actions (MacMahon et al. 2006). In both cases, a sophisticated dialogue model can support the inference processes by filling in missing information with

respect to both the implemented spatial model, and the real world in which the interaction takes place.

The present Wizard-of-Oz study was purposively designed to assume more mismatches than would normally be the case using any sophisticated spatial language understanding system. Nevertheless, the need for conceptual clarification questions will remain, particularly with increasing spatiotemporal complexity. Such procedures are well known also from human-human interaction (which may be assumed as a "gold standard" for our research), e.g., Filipi & Wales (2004). In the present study, the clarification attempts by the robot worked best for the discourse flow when they could be integrated into the user's current mental representation of the spatial as well as the discourse situation. In other cases, clarification questions could induce spatiotemporal distortions not encountered in our previous monological experiments (Shi & Tenbrink *forthc.*), thus complicating the dialogue rather than enhancing it.

Robotic requests that include a new starting point, such as "When I am left of the room with the photocopying machine, what do I do?" were taken up easily by the users especially in cases of earlier confusion. To generalize this idea, it is important that the robot informs the user about its current state of knowledge in as much detail as possible, and suggests a solution concerning how to proceed further. This will be specifically helpful in the case of spatiotemporal sequencing confusions. Also, it is important that the robot acknowledges what it has understood so far, to let the user know where exactly there is an information gap that needs to be filled in, and to align the spatiotemporal concepts that the interactants are currently referring to. These results are related to Rieser & Moore's (2005) finding that it is better for systems to ask for confirmation of a hypothesis than to merely signal non-understanding.

In general, our brief investigation of a situated dialogic interaction in which a robot's reactions were simulated shows that requesting clarification about spatial representations is a non-trivial endeavour in which even slight deviations in timing or in confirming common ground may lead to severe distortions (see also Stoia et al. 2006). With a real robotic system, speech recognition problems will complicate the situation considerably (Moratz & Tenbrink 2003), although more standard clarification procedures (Purver et al.

2003, Schlangen 2004) are then applicable to cover some of the problems.

#### 4 Improvement of the dialogue model

Regarding the results of our analysis, the dialogue model used as motivation for the empirical studies (cf. section 2.2) needs to be extended. This concerns, in particular, an improvement of the clarification procedures, the amount of feedback provided by the robot, and a more precise matching process between system knowledge and the linguistic input by the user. Specifically, the precise discourse history is important since specific requests providing information about successfully integrated knowledge are more useful than generic clarification questions, as motivated above. Moreover, the robot's internal map represented as a Conceptual Route Graph and the robot's current position on the map should be used for informing the user in detail about current disparities, in order to classify various requests, and to make precise suggestions (see below). In the former version, this information was only used to detect mismatches, not to inform the user within the clarification sub-dialogues. To achieve an effective and natural dialogue with users, the dialogue model needs to take account of information from both dialogic and internal sources. Consequently, the first extension of the dialogue model augments it by integrating the dialogue history as well as the internal map with the robot's current position (denoted as  $[H, M]$ ). The CONversational Roles model is a generic situation-independent dialogue model. Dialogue models based on the COR model cover discourse patterns that are independent of the dialogue context. By integrating the dialogue history as a parameter in the extended dialogue model we add a crucial element from the well-known information state approach (Traum & Larsson 2003) into the dialogue modelling process. As a result the model benefits from both approaches: the flexibility of the information state approach and the well defined structure of the COR based modeling approach.

With respect to the mapping of user utterances to the robot's internal map, the general utterance "this does not match with my internal map" did not seem to be helpful for the users but rather caused confusion (cf. D3). Precise suggestions such as "Is this the first possibility to turn?" seemed more promising (cf. D1). In our improved model, we

substitute the three simple dialogue acts, *request*, *assert* and *suggest* (see Fig. 1 above) by subdialogues. Each subdialogue uses the discourse history and the internal map representation to support detailed classifications. Figure 2 represents the sample subdialogue *request(robot,user)*. First, the robot acknowledges the part of the instruction that it has understood, based on [H, M]. The user can react by rejecting this account and providing a further instruction, in which case the robot does not formulate the request in the intended way. However, if the user does not react or reacts by accepting the robot's description, the robot continues by requesting information about entities, boundaries, or orientations, depending on the current requirements, in a way that is aligned to the users' descriptions as much as possible (using the dialogue history). The dialogue will then continue with the user providing the missing information.

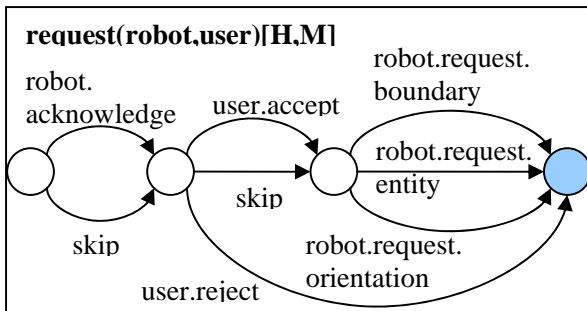


Figure 2 'Request' subdialogue

## 5 Conclusion and Outlook

We presented a detailed qualitative analysis of a Wizard-of-Oz study specifically tailored to the intended functionalities of the robotic wheelchair Rolland, employing the first version of our dialogue model. Results show that the model is successful in encouraging the user to provide missing information and to use a suitable level of granularity. However, clarification questions from the robot need to be formulated and placed with specific care, as even slight confusions and temporal misplacements of the robot's utterances can lead to severe communication problems and distortions of the user's spatiotemporal representation. Our proposed solution is to let the robot inform the user about its internal state of knowledge in as much detail as possible, and to formulate requests and suggestions in a way that is aligned to the user's

descriptions. The next step in our iterative approach is to test this revised model empirically.

The construction of dialogue models is the first step towards the development of dialogue systems based on empirical findings. We are now developing a general approach to specify straightforwardly Recursive Transition Networks in a formal specification language, using the model-checker technique to analyse features, complexity and coverage of dialogue models. Then, dialogue models will be constructed from empirical data by extracting the discourse patterns from annotated dialogues, and analysing the relations between discourse patterns and dialogue models. This procedure will enable us to assert how many dialogues fall into a given dialogue model, which may serve as a basis for evaluating a dialogue's success and efficiency and comparing various instances of dialogue systematically. This approach also supports the mechanical comparison of dialogue models and can thus be used in the dialogue model evaluation process in future iterations.

## References

- Bugmann G., Klein E., Lauria S. and Kyriacou T. 2004. Corpus-Based Robotics: A Route Instruction Example. In *Proc. IAS-8*, Amsterdam, pp. 96-103.
- Clark, B., O. Lemon, A. Gruenstein, E. Owen Bratt, J. Fry, S. Peters, H. Pon-Barry, K. Schultz, Z. Thomsen-Gray, and P. Treeratpituk. 2005. A General Purpose Architecture for Intelligent Tutoring Systems. In J. C. J. van Kuppevelt, L. Dybkjær and N.O. Bernsen (eds.), *Advances in Natural Multimodal Dialogue Systems*. Berlin, Heidelberg: Springer.
- Clark, H.H. and D. Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition*, 22:1-39.
- Denis, M., F. Pazzaglia, C. Cornoldi, and L. Bertolo. 1999. Spatial discourse and navigation: an analysis of route directions in the city of Venice. *Applied Cognitive Psychology* Vol. 13/2, pp. 145 - 174.
- Filipi, A. and R. Wales. 2004. Perspective-taking and perspective-shifting as socially situated and collaborative actions. *Journal of Pragmatics, Volume 36, Issue 10*, 1851-1884.
- Fischer, K. 2003. Linguistic Methods for Investigating Concepts in Use. In Stolz, T. and Kolbe, K. (eds.), *Methodologie in der Linguistik*. FfM: Lang, 39-62.
- Gieselmann, P. and A.Waibel 2005. What makes human-robot dialogues struggle? In *Proc. DIALOR*.

- Gryl, A., B. Moulin and D. Kettani. 2002. A conceptual model for representing verbal expressions used in route descriptions. In K. Coventry and P. Olivier (eds.), *Spatial Language: Cognitive and Computational Perspectives*. Dordrecht: Kluwer, pp. 19-42.
- Hois, J., M. Wünnel, J.A. Bateman, and T. Röfer. 2007. Dialog-Based 3D-Image Recognition Using a Domain Ontology. In T. Barkowsky, M. Knauff, G. Ligozat, and D. Montello (eds.), *Spatial Cognition V: Reasoning, Action, Interaction*. Berlin: Springer.
- Klippel, A., T. Tenbrink, and D. Montello (in press). The Role of Structure and Function in the Conceptualization of Directions. In E. van der Zee and M. Vulchanova (eds.), *Motion Encoding in Language and Space*. Oxford University Press.
- Krieg-Brückner, B. and Shi, H. 2006. Orientation Calculi and Route Graphs: Towards Semantic Representations for Route Descriptions. In Raubal, M., Miller, H.J., Frank, A.U., & Goodchild, M.F., *Proc. GIScience 2006, Münster*. Berlin: Springer, pp 234–250.
- Kruijff, G.-J., H. Zender, P. Jensfelt, and H.I. Christensen. 2007. Situated Dialogue and Spatial Organization: What, Where... and Why? *International Journal of Advanced Robotic Systems, Special Issue on Human and Robot Interactive Communication* 4, 2.
- Kruijff-Korbayová, I., E. Karagjosova, and S. Larsson. 2002. Enhancing collaboration with conditional responses in information-seeking dialogues. In Bos, Foster and Matheson (eds): *EDIALOG 2002*, 4-6 September 2002, Edinburgh, UK, pp. 93-100.
- Lankenau, A., and Röfer, T. (2000). The Role of Shared Control in Service Robots - The Bremen Autonomous Wheelchair as an Example. In: Röfer, T., Lankenau, A., Moratz, R. (Eds.): *Service Robotics - Applications and Safety Issues in an Emerging Market. Workshop Notes, ECAI 2000*. 27-31.
- Lemon, O., A. Bracy, A. Gruenstein, and S. Peters. 2003. An information state approach in a multi-modal dialogue system for human-robot conversation. In P. Kühnlein, H. Rieser and H. Zeevat (eds.): *Perspectives on Dialogue in the New Millennium*. Amsterdam: Benjamins, pp 229-242.
- MacMahon, M., B. Stankiewicz, and B. Kuipers. 2006. Walk the Talk: Connecting Language, Knowledge, and Action in Route Instructions. In *Proc. AAAI-2006*, Boston, MA.
- Moratz, R. and Tenbrink, T. 2003. Instruction modes for joint spatial reference between naive users and a mobile robot. *Proc. RISSP 2003*, Special Session on New Methods in Human Robot Interaction, October 8-13, 2003, Changsha, Hunan, China.
- Purver, M., Ginzburg, J. & Healey, P. 2003. *On the means for clarification in dialogue*. In Smith, R. and van Kuppevelt, J. (eds.), *Current and New Directions in Discourse and Dialogue*. Kluwer, pp. 235-255.
- Rieser, V. and J.D. Moore. 2005. Implications for Generating Clarification Requests in Task-oriented Dialogues. *Proceedings of the 43rd Annual Meeting of the ACL*, pp. 239–246, Ann Arbor.
- Ross, R.J., H. Shi, T. Vierhuff, B. Krieg-Brückner, and J.A. Bateman. 2005. Towards Dialogue Based Shared Control of Navigating Robots. In Freksa, C., M. Knauff, B. Krieg-Brückner, B. Nebel, and T. Barkowsky (eds.), *Spatial Cognition IV: Reasoning, Action, Interaction*. Berlin: Springer, pp. 479-500.
- Schlangen, D. 2004. Causes and strategies for requesting clarification in dialogue. In *Proc. SIGdial04*.
- Selting, M. 2000. The construction of units in conversational talk. *Language in Society* 29, 477–517.
- Shi, H. and T. Tenbrink (forthc.) Telling Rolland where to go: HRI dialogues on route navigation. In K. Coventry, T. Tenbrink, and J.A. Bateman (eds.), *Spatial Language and Dialogue*. Oxford University Press.
- Sitter, S. and A. Stein. 1992. Modelling the Illocutionary Aspects of Information-Seeking Dialogues. *Information Processing and Management* 28, 124-135.
- Spexard, T., S. Li, B. Wrede, J. Fritsch, G. Sagerer, O. Booij, Z. Zivkovic, B. Terwijn, and B. Kröse. 2006. BIRON, where are you? Enabling a robot to learn new places in a real home environment by integrating spoken dialog and visual localization. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.
- Stoia, L., D.K. Byron, D. Shockley, E. Fosler-Lussier. 2006. Sentence Planning for Realtime Navigational Instruction. *Proc. HLT-NAACL 2006*, 157-160.
- Tenbrink, T. 2007. *Space, time, and the use of language: An investigation of relationships*. Berlin: Mouton de Gruyter.
- Traum, D. and Larsson, S. 2003. The Information State based Approach to Dialogue Management. In R. Smith and J. van Kuppevelt (eds.), *Current and New Directions in Discourse and Dialogue*. Dordrecht: Kluwer, pp. 325-353.
- Tversky, B. and P. Lee. 1998. How Space Structures Language. In: C. Freksa, C. Habel & K.F. Wender (eds.), *Spatial Cognition. An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*, pp. 157-175.