Building an Empirically Founded Dialogue System

Thora Tenbrink, Shi Hui, Robert Ross, Elena Andonova,
Juliana Goschler, John Bateman

SFB/TR 8
SPATIAL COGNITION

Contact Address:

Dr. Thomas Barkowsky
SFB/TR 8   Tel +49-421-218-64233
Universität Bremen   Fax +49-421-218-64239
P.O.Box 330 440   barkowsky@sfbtr8.uni-bremen.de
28334 Bremen, Germany   www.sfbtr8.spatial-cognition.de

Building an empirically founded dialogue system

**Tenbrink, Thora, Shi Hui, Robert Ross, Elena Andonova,
Juliana Goschler, and John Bateman**

Abstract

The major aim in our project is to enable flexible dialogue management for intuitive spatial communication. This requires the dynamic negotiation of spatial reference frames, spatial perspectives and discourse strategies. Speakers manage such negotiation by the fine-scale selection and alignment of linguistic forms sensitive to dialogue history, user-group defined norms, and spatial context. Such mechanisms are empirically derived and implemented within our dialogue system for natural interaction in spatially-embedded tasks. In an iterative and cyclic approach, we combine empirical methods of discourse-analytic and psycholinguistic investigation of dialogue in the spatial domain with formal dialogue modelling and the specification and implementation of a dialogue system. From an empirical point of view, the particular challenge lies in doing adequate justice to the conflicting tensions of, first, eliciting free language production from experimental participants for realistic interactional data, second, enforcing sufficient experimental control to warrant statistically validated claims, and third, evaluating the current state of the system with untrained users interacting intuitively with the system, based on adequately projected discourse aims and tasks. In this paper we introduce our general approach, describe each of the involved research directions and their interrelation, discuss challenges, and highlight ensuing synergies, including the requirement of operationalizing annotation categories precisely.

1 Introduction

What features and functionalities does a natural language based dialogue system need in order to render communication with an untrained human fluent, efficient, and intuitively natural? In our project, we address this question with respect to a limited but pervasive and well-researched domain, namely, space: natural language employed in relation to settings necessitating reference to spatial concepts, entities, and relationships.

Dialogic spatial interaction, a major area of importance for spatially-aware systems, remains under-represented in the literature. Specifically, it is an open question how speakers' choices of conceptual reference systems and their linguistic representations are influenced by the discourse history and by the interlocutor's feedback. For example, our own earlier findings demonstrate the strong impact of initial utterances (interactive alignment via priming) on subsequent spatial descriptions in terms of reference frame choice. This issue becomes increasingly crucial in the area of human-robot interaction, as it is by now well-established that speakers react intensively to the requirements of their artificial interaction partner, both with respect to linguistic choices (Amalberti et al., 1993) and high-level decisions (Hinds et al., 2004). Even small changes in the experimental setting, including the robot's reactions, may be crucial in this regard (Moratz & Tenbrink, 2006; Fischer, 2006), along with users' preconceived mental models and expectations that are equally decisive for users' conceptualization of the dialogue and their ensuing linguistic reactions (Clark, 1999, Andonova, 2006). Schober and Brennan (2003) extensively discuss a broad range of research addressing various aspects of dialogic interaction, concluding that the processes involved are much more complex than previously assumed.

Existing evaluations that have been carried out in human-robot interaction (HRI) without restricting in advance the language that may be adopted by users have shown that systems can do very badly, simply because the actual language used lies outside of that supported (Thrun 2004). This situation is made worse when users gain an overly-positive impression of the system's capabilities – for example, by canning more sophisticated responses than a system can actually meaningfully produce. It is therefore essential for HRI to be based on realistic assessments both of what language users will produce and of the consequences that the robot's dialogue contributions

bring for their users' subsequent language choices. The importance of interdisciplinary, combined-method approaches to this problem is discussed further in, for example, Burke et al. (2004).

To handle such known problems and difficulties we combine several methods to enable the development of flexible and adaptive dialogic interaction with intelligent systems, employing an iterative development process. As an application scenario, we are assuming a mobile autonomous service robot for home usage, in the baseline case simply with the purpose of making life easier for non-handicapped users, targeting specific applications for physically handicapped as well as elderly users in later progress. The service robot should be able to understand spoken and written natural language input, react verbally and behaviorally to instructions, ask clarification questions, and navigate autonomously. In a particular situation, the robot may have partial knowledge about the environment but not about the exact location of particular objects like the coffee machine, while the human knows all relevant details of the actual environment. Conversely, a situation may occur in which the human user has less knowledge about the environment than the robot. Both cases necessitate the negotiation of spatial relationships and routes. Then, communication problems may arise, for instance, due to mismatches in the knowledge of the interlocutors as will be described below.

We address users' spontaneous interactional strategies for avoiding communication problems during spatial tasks; we also investigate how clarification dialogues can be initiated by agents in an effective way in cases of communication failure. Related to these procedures in terms of collaborative negotiation, we focus on particular areas where the basic premises of interactive alignment – i.e., that it operates by 'priming' in the psychological sense – is not appropriate, even though an effect similar to alignment is achieved; this relates to research in areas such as *accommodation* (Giles & Coupland, 1991) and *resonance* (Sakita, 2006). Both of these accounts assume a fairly high level of awareness by the speakers concerning strategies of convergence and parallelism, contrasting (implicitly) with the mechanistic account proposed by Pickering and Garrod (2004). Beyond gaining insights on these processes, our basic assumption is that negotiative alignment should be made applicable for computational dialogic systems, including our larger-scale computational grammars for genera-

tion and analysis and the formal ontologies that we employ to mediate between natural language components and domain knowledge.

In the following section, we describe our approach in general terms. Then we address each of the involved directions of research in more detail, along with an assessment of their relative significance for each other. In Section 4 we discuss and elaborate some of the particular challenges met in the process of our ongoing project work.

2 General approach

In order to integrate implicit and explicit negotiation mechanisms in computational dialogue systems we draw on a complementary set of research methods for constructing models of user-adaptivity. Drawing on our earlier findings resulting from extensive use of Wizard of Oz dialogues and fixed protocols, our current focus is on the availability of our dialogue system in order to explore and refine the dialogic mechanisms. Our approach is both cyclic and iterative, as illustrated in Figure 1; it involves empirical investigation and analysis, modelling processes, and implementation in the system.
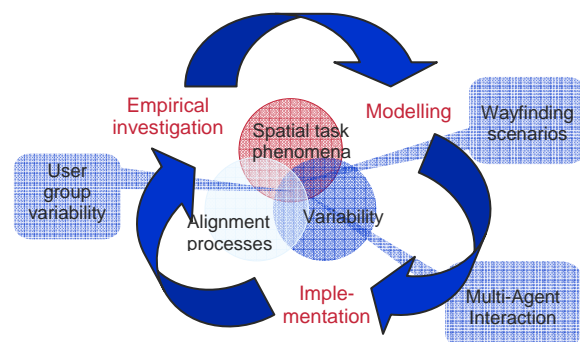

Figure 1. Iterative design

We target a careful coordination of established psycholinguistic experimentation on the one hand and qualitative empirical discourse analysis of 'freely' produced dialogic contributions in a spatial interaction on the other as the method of choice for reliably revealing the fine-details of negotiative alignment. To support this, we employ both human-human baseline-establishing experiments and genuine HRI and human-system interactions. For the latter the dialogue system is progressively augmented with automatic adaptation according to user model as the empirical results are transferred. This requires the development of a generic computational approach to adapting the dialogue system according to users' dialogic behaviour.

In particular, we isolate linguistic negotiation phenomena that fall within the targeted competence of our dialogue system and correlate their occurrence or non-occurence with controlled situational and dialogic conditions. The experimental setups address the levels of linguistic/spatial representation where negotiation is hypothesized to operate: lexical, grammatical, semantic, spatial perspective, and dialogue strategy, including choice of granularity level (reference to the goal location or a description of the path towards it), reference system (absolute, relative, or intrinsic), perspective (user-centered or robot-centered), relata (which objects can reliably be referred to, and in what ways?), and reference axis – all of which may be interdependent and are known to be influenced by further discourse factors other than simple priming-based alignment (cf. Tenbrink, 2007). We also vary the spatial configurations within which interaction is undertaken in order to pinpoint when speakers align with each other, when they predominantly align with the world, and when there is some combination of these effects.

The linguistic data are transcribed, annotated and analysed focusing on the points of variation, highlighting systematic patterns of language usage and conceptual choices in relation to the spatial situation and discourse task, and making precise the consequences of alignment processes initiated by the dialogue system on users' subsequent utterances. Crucially, annotation is geared towards the particular fine grammatical variation that our dialogue system can control. The data result in a resource pool of annotated corpora. From the results, hypotheses are formulated to bring out consequences that feed directly into the computational model. In the following we spell out how this general approach is tackled by the diverse methodologies involved.

3 Experiences

3.1 Psycholinguistic concerns

In the area of psycholinguistics, researchers aim at the identification of experimental designs that allow sufficient control of human linguistic behavior to warrant statistically validated claims. Empirical elicitation concerns a particular kind of variability in human responses in relation to a controlled setting. This setting is chosen to allow for a theoretical assessment of the significance of this variability, addressing an open research question that has not been sufficiently resolved in the available literature so far. It is a particular challenge in this area to balance between the need for experimental control on the one hand, and the need for generalizability, theoretical significance, and ecological validity on the other hand. Even within the psycholinguistics community itself, this challenge is well-known and much discussed: a particular asset of any publication in this area concerns its success in establishing how the specific results gained in a controlled experimental setting relate to the more global psychological/psycholinguistic or cognitive processes active in any kind of linguistic (or related) activity, at best in relation to everyday life.

That said, the application of this approach in our interdisciplinary work nevertheless raises additional questions not typically covered in this type of research. In addition to the need to generalize reliably, the need arises to identify precisely those psycholinguistic processes that are relevant for, on the one hand, the development of the dialogue system as such, and on the other hand, the particular settings and application areas for which the dialogue system is being developed. However, as yet there is only very little research concerning the psycholinguistic processes involved in human-computer interaction of any kind. Branigan et al. (2003) provide relevant evidence concerning syntactic alignment processes in human interaction with a computer. These findings relate to Zoltan-Ford's (1991) proposal to influence users' linguistic choices by particular types of computer output. However, precisely how these alignment processes work in the spatial domain has not been established yet. Furthermore, even in the area of human-human dialogue, spatial negotiation processes have not been investigated sufficiently so far, with some exceptions (e.g., Coventry et al., in press). Therefore, from a psycholinguistic point of view, the most urgent aim at present is to identify processes such as alignment with respect to a setting that resembles the application areas for our dialogue system in general, but involves humans engaged in a kind of dialogue that is not perceived as simulating the interaction with a system.

To provide an example, one recent study targets the following issue. When negotiating a route to be travelled by an agent with respect to a two-dimensional map, two kinds of perspective need to be distinguished (Taylor and Tversky, 1996): survey (looking at the map from "outside" the scene) versus route perspective (using the perspective of the route-travelling agent). This

distinction becomes relevant for our purposes when the agent travelling the route is controlled by a computer informed by the dialogue system, for example, in order to demonstrate a service robot's future path or to visualize a route in reaction to a request by a human. In this respect we need to establish to what extent users react to the previous utterances heard by an interlocutor (which, ultimately, is intended to be the dialogue system). In other words, based on the findings described above (in addition to our own findings on human alignment in spatial contexts, Coventry et al., subm.) we expect users to be influenced by the dialogue system's contributions; our interest now lies in the degree to which such an influence is reliable and utilizable for enhancing successful communication. The psycholinguistic approach contributes valuable findings in this regard by adopting a restricted experimental setting involving one single direction change at a time. Two human interlocutors take turns to describe this change. Our study (the analysis of which is currently underway) uses a confederate paradigm in which one of the human interlocutors talks according to a carefully designed script, with a controlled sequence of route and survey perspectives expressed in a pragmatically natural way. The dependent variables then concern the naïve participant's linguistic contributions, i.e., the degree to which the speakers are influenced by their interlocutor's choice of perspective.

3.2 Discourse-analytic concerns

In the area of discourse analysis, major aims concern the identification of linguistic patterns in discourse and their relationship to such factors as the text type, the general setting, and the discourse history. Thus, from a discourse analytic perspective it is natural to look at the peculiarities of HRI, and to investigate how diverse factors of the spatial setting contribute to speakers' linguistic choices. Crucially, however, such choices are not identified in terms of fine-grained restricted choices as in the case of psycholinguistic-type studies, but in terms of a more general view of "what happens" in discourse, establishing a more thorough understanding of how speakers build up a text or dialogue given the current situation. Such studies are well suited for outlining dialogic negotiation principles and patterns that dialogue systems need to account for, and for formulating precise hypotheses that can then be tested in accurately controlled and restricted studies.

In our earlier work we identified a systematic and fine-grained account of the speakers' repertoire, principles, and patterns in using spatial language with regard to diverse spatial settings and tasks (Tenbrink, 2007). Results of this work have been integrated, on the one hand, in the linguistic ontology adopted in our dialogue system, and on the other hand, in the development of the required range of vocabulary and grammar, as well as a number of pragmatic principles accounted for by the system. The next steps in this regard concern the particular *dialogic* processes involved in spatial interaction. For this purpose, we collect dialogic language data in naturalistic scenarios that are kept as close as possible to the application area of the dialogue system. The dialogues are investigated with respect to systematic patterns such as the mutual negotiation of spatial reference frames, spatial perspectives, scale and granularity choices, and corresponding discourse strategies adopted by the speakers.

As an example, consider the perspective choice problem already addressed in the previous subsection. From a discourse-analytic point of view, interesting research questions concern, on the one hand, the range of linguistic choices and markers of spatial perspective, and on the other hand, the dialogic developments that eventually lead to a particular choice of perspective, going beyond immediate subconscious processes of priming and alignment. Therefore, the setting we have adopted for collecting results with respect to our perspective problem allows for more freedom of choice than that possible within the psycholinguistic-type study described above. Our recent study (Goschler et al., 2008) involves a schematic map and two naïve participants who were asked to imagine being situated in the environment. One participant was asked to give instructions for navigating towards a pre-defined goal that their partner couldn't see. The other participant was asked to imagine sitting in the wheelchair and navigating towards the goal according to their partner's instructions. Given this setting, the participants were allowed to use their own linguistic strategies. Accordingly, the heterogeneity in the data is considerable. Nevertheless, a number of relevant patterns emerge that provide a useful basis for further research both in psycholinguistics and in discourse analysis, as well as for the development of the dialogue system.

For instance, a range of linguistic markers of perspective typical of this setting could be identified, which is essential for an operationalization

of coding categories when analyzing natural language data (see Section 4) as well as for the vocabulary extension of the dialogue system. Furthermore, the analysis reveals how particular conceptual patterns are linguistically reflected, such as perspective choices and shifts, or variations of granularity levels for the benefit of incremental navigation instruction.

A further benefit of the experimental design lies in the combination of approaches adopted in our project, concerning the option of adding in research questions from a psycholinguistic perspective. While free language production data are generally not particularly suitable for statistical analysis, the broad interest in the area of psycholinguistic studies (as outlined above) in achieving ecological validity results in substantial experience in identifying particular general factors (abstracting from linguistic details) that may be addressed with respect to less controlled data sources such as the present one. Here, we carried out analyses concerning the relative dominance of use of the two perspectives across dyads and speakers, and the degree to which there was coordination between instructors and instructees in individual dyads in terms of preferences for a certain perspective. The former analysis revealed that, although in the corpus as a whole, route perspective utterances were far more numerous, if perspective choices are examined within dyads, we find a considerable amount of variation, and no clear dominance of one of the two perspectives. The latter analysis further highlighted the existence of a high degree of coordination between interlocutors in the individual dyads in terms of the number of their dialogic contributions - the more the instructor spoke, the more the instructee said as well, and vice versa (Goschler et al., 2008).

Beyond results of this kind, that are basically motivated from prior research, a further opportunity to benefit from the results consists of a direct (though post-hoc) statistical validation of a number of hypotheses motivated from the discourse-analytic point of view. The argument then goes as follows. If a particular process is indeed active, it should lead to specific patterns in the linguistic data (at best, binary distinctions) that can be distilled from the corpus in a targeted way. In our case, we identified a number of hypotheses for how and why speakers undertake perspective shifts in the middle of a conversation. These hypotheses could then be directly addressed by statistical tests, albeit to a limited degree, motivating more controlled research for the future.

With respect to dialogic interaction the linguistic and structural patterns of interaction themselves are also of great importance to our understanding of natural dialogue. Within such structural modelling, the aim is to identify, specify, and formalize systematic patterns in dialogue, such as particular kinds of dialogue acts. Several schemes for the abstract description of dialogue processes have been produced (e.g., DAMSL, Allen & Core, 1997), and these can in turn form the atomic elements of larger scaled structural accounts (Sitter & Stein 96) which allow us to derive patterns of language broadly independent of any specific theory of dialogue competence (Shi et al., subm.).

A prerequisite for such analysis lies in the annotation of natural dialogue data based on a suitable coding scheme, and extracting the relevant patterns automatically. Human-human dialogues exhibit a range of interpretation and clarification strategies that ensure that interaction usually proceeds smoothly, with clarifications being introduced as necessary and with considerable precision. We specify these processes through the reuse of the corpora collected for other aspects of our discourse-analytic research as just described. Thus, our by now extensive corpus of free production data collected in spatial settings across a range of types of interaction (between humans, or between humans and – robotic or other – systems) allows for a multi-level analysis of dialogic effects both from a dialogue structure modelling perspective, and in terms of extracting systematic linguistic patterns concerning how humans represent spatial relationships, and how they interact with robots under varying circumstances (Moratz & Tenbrink, 2006; Shi & Tenbrink, 2009; Vorwerg & Tenbrink, 2007).

3.3 Computational Implementation

The operationalization of empirical results obtained from our psycholinguistic and discourse analysis work requires semantically consistent computational formalisms which are rich enough to capture the details of these models, yet tractable in application within our targeted robotic systems. Indeed, the implementation of our applied dialogue systems itself involves a complex range of computational issues (see Ross et al., 2005). Here we focus particularly on how our development draws directly on the results of our empirical study methodologies.

For the application of structural dialogue models, we employ a formal method based approach known as *Communicating Sequential Processes*

(CSP) (Hoare, 1985) both for the representation of the derived model, and for its deployment within computational systems. This allows for the straightforward iterative development of dialogue models as well as for a precise comparison of distinct dialogue models at varying levels of abstraction; moreover, it also offers suitable means for checking that specifications conform to the dialogue models required and their desired properties.

A major challenge targeted specifically in our project lies in the translation of verbal spatial descriptions to the resources of spatial knowledge available to the system. Since a human's perception and verbalization of the environment differs substantially from a robot's implemented map or knowledge derived from perceptual functionalities, no direct mapping is possible. This leads to a broad range of potential mismatches. Such complexities motivate our principled analysis of naturalistic language phenomena and their inclusion into the dialogue system. This is captured in our approach of applying empirical findings to the development of fine-grained linguistic semantics models which have been developed on the basis of detailed linguistic ontologies of spatial language use (Bateman et al., subm.). In turn these models have been cast within appropriate grammars of language analysis and production based on the formalisms of Categorial Combinatorial Grammar (Steedman 2001) and Systemic-Functional Grammar respectively (Halliday & Matthiesen 2004).

Such computational grammars and semantics provide detailed accounts of language use, but must be further extended to account for use in context. This concerns, on the one hand, our findings on situated language use in controlled and data intensive experiments, but also embodied models of spatial meaning, and the mechanisms which support spatial reasoning over complex spatial environments: for example, spatial calculi such as the Region Connection Calculus (Randell et al., 1992) or the Route Graph (Krieg-Brückner & Shi, 2006). Furthermore, the identification of discourse referents typically used as landmarks within route instructions can depend on a range of spatial factors such as visual saliency, proximity, or accessibility relations. Also, the mapping process can involve the application of non-physical context to enrich the surface information provided. While the application of contextual information in the transformation process is desirable, a suitable clarification process as specified in the previous subsection remains inevitable for those situations where context application alone fails to resolve underspecification or uncertainty.

Following the implementation of the diverse modules and procedures necessary for the dialogue system, the obvious empirical step in this area lies in obtaining suitable evaluation data based on the confrontation of naïve users with the system. For this purpose, we carry out studies with a limited amount of participants at intermediate steps of the system development procedure, targeting an iterative process such as that suggested by Moratz & Tenbrink (2006). This procedure reduces the cases of communicative failure that are due to insufficient vocabulary or grammatical coverage to a minimum, while ensuring that the psycholinguistic and discourse-analytic results obtained from human-human interaction studies are transferred successfully to the system's functionalities. Gradually, the same analytic procedures as developed for human dialogues can be adapted for human-system interaction data, allowing for a direct comparison in those cases where one particular experimental design can be successfully adopted for both types of interaction.

4   Challenges

With respect to the diverse procedures adopted in our project as outlined so far, we encounter a range of challenges that are not typically a subject of targeted investigation. On the one hand, there are the (by now almost notorious) problems of identifying suitable publication procedures and outlets for interdisciplinary research, due to the diversified goals addressed simultaneously in each of the areas involved. Naturally, for example, a "purely" psycholinguistically motivated experimental design can neglect issues such as the transferability of a particular scenario to a HRI setting. On the other hand, the adjustment of various methodological approaches towards one aligned approach presupposes the outline of a suitable scenario such as that indicated in Section 1 above, which – in the context of fundamental scientific (rather than applied technological) research – may not be conceived of as self-evident.

Also, model operationalization and application brings significant challenges for our computational modeling work, for example, the combination of standard dialogue management approaches (e.g., information state) and formal method based dialogue control; integrating spatial models; and the incorporation of empirically

valid models of spatial language within our linguistic resources and processing frameworks. While each of these issues are in themselves technically difficult, and certainly beyond the scope of this paper to detail, the challenge is also to draw on our empirical results in a methodological way such that we both capture our psycholinguistic and discourse analytical results, and yet provide suitable feedback into such studies, thus actively combining the research objectives and methodologies of what are often seen as disparate research fields.

Furthermore, particular steps of the analysis procedure may turn out to be challenging. Here, however, we view the challenge as a particular supportive feature of the interdisciplinary approach, since the requirement of modelling formally, or operationalizing annotation procedures for the purposes of automatic processing, leads the analyst towards a more precise specification of the observed phenomena than encountered elsewhere. With respect to annotation, one of our most crucial concerns is thus the identification of well-defined criteria suitable for a type of analysis that aims at a maximum level of operationalizability and objectivity. With respect to our example explored above concerning perspective choice, this type of analysis starts from the identification of particular linguistic forms that signal a particular kind of perspective. But since trained human annotators seldom encounter problems in identifying the underlying perspective (or the lack of differentiation of perspective) even in those cases in which no such direct mapping is available, there must be further identifiable criteria obtainable from the discourse and situational context that are not as readily observable. For much earlier research, it has been sufficient to simply annotate the result of such intuitive differentiation by the analyst. For our purposes, we aim at capturing precisely those factors that lead the analyst towards the diverse distinctions. Naturally, any success concerning operationalization will directly result in an enhanced level of inter-coder agreement. Clearly, a full operationalization requires substantial iterative and fine-grained analysis and even then does not always lead to complete success. But even at an intermediate level, considerable benefit lies in identifying particular factors that are easy to specify precisely and (nearly) objectively, rather than trying to cover the full range of interesting aspects in the data traded against a reliable, accurately specified coding scheme.

As an example, the DAMSL specification (Allen & Core, 1997) suggests that a decision be made concerning whether the speaker is "trying to change the belief of the addressee". Such a definition may be intuitively appealing and helpful in the annotation process according to the specified scheme, but no direct operationalization procedure is offered. This kind of specification will hardly be derivable directly from the data in any generalized way, given the broad range of linguistic options available to speakers to formulate their underlying intentions. Nevertheless, language use is not a random process, and particularly with respect to a well-researched domain such as spatial language it is possible to identify systematic patterns as indicated above. Thus, our endeavour at this point lies in formulating and defining precisely what it means in our context to "try to change the addressee's belief". The formal modelling of dialogue acts then allows for the selection of an appropriate formulation from the available resource pool of possible linguistic specifications, supporting intuitive and empirically founded communication between untrained users and a natural language based dialogue system.

5 Conclusion

The interdisciplinary and combined-methods approach in our project entails a range of challenges that, when overcome, result (among other effects) in a well-founded basis for fine-grained reliable analysis of natural language data in spatial settings. We hope with this contribution to encourage cross-community communication for the benefit of a more thorough understanding of natural dialogue procedures, ultimately enabling intuitive and flexible interaction between humans and artificial systems of any kind.

References

Allen, James and Mark Core. 1997. *Draft of DAMSL: Dialog Act Markup in Several Layers*. Manuscript, http://www.cs.rochester.edu/research/speech/damsl/RevisedManual/.

Amalberti, R., N. Carbonell, and P. Falzon (1993). User Representations of Computer Systems in Human-Computer speech interaction. *International Journal of Man-Machine Studies,* 38:547-566.

Andonova, Elena. 2006. On Changing Mental Models of a Wheelchair Robot. In Kerstin Fischer (ed.), *Proceedings of the Workshop on 'How People Talk to Computers, Robots, and Other Artificial Communication Partners'*, Hansewissenschaftskolleg,

Delmenhorst, April 21-23, 2006, pp. 131-139. SFB/TR8 Report 010-09/2006.

Bateman, John, Joana Hois, Thora Tenbrink, and Robert Ross (subm). A Linguistic Ontology of Space for Natural Language Processing.

Branigan, Holly, Martin Pickering, J. Pearson, J. McLean, and C. Nass. 2003. Syntactic alignment between computers and people: The role of belief about mental states. In *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, pp. 186–191.

Burke, J.L., Murphy, R.R., Rogers, E., Lumelsky, V.J., and Scholtz, J. 2004. Final report for the DARPA/NSF interdisciplinary study on human-robot interaction. *IEEE Transactions on Systems, Man and Cybernetics, Part C*, Vol. 34 No.2, pp. 103-112.

Clark, Herbert H. 1999. How do real people communicate with virtual partners? In *Proceedings of AAAI-99 Fall Symposium, Psychological Models of Communicatin in Collaborative Systems,* November 5-7th , 1999, North Falmouth, MA. Menlo Park, Calif.: AAAI Press; Cambridge, Mass.: MIT Press.

Coventry, Kenny R., Elena Andonova, and Thora Tenbrink (subm.). Aligning with what we hear and what we see: Reference frame use and verbal and visual context.

Coventry, Kenny, Thora Tenbrink, and John Bateman (eds., in press). *Spatial Language and Dialogue.* Oxford University Press.

Fischer, Kerstin. 2006. *What Computer Talk Is and Is not: Human-Computer Conversation as Intercultural Communication*, volume 17 of *Linguistics - Computational Linguistics*. Saarbrücken: AQ.

Giles, H. and N. Coupland. 1991. *Language: Contexts and Consequences*. Keynes: Open University Press.

Goschler, Juliana, Elena Andonova, and Robert Ross. 2008. Perspective Use and Perspective Shift in Spatial Dialogue. In Christian Freksa, Nora Newcombe, Peter Gärdenfors, and Stefan Wölfl (Eds.), *Spatial Cognition VI: Learning, Reasoning, and Talking about Space*. Berlin: Springer, pp. 250-265.

Halliday, Michael A.K. and Christian Matthiessen. 2004. *An Introduction to Functional Grammar.* Edward Arnold, London, 3rd edition.

Hinds, Pamela J., Teresa L. Roberts, and Hank Jones. 2004. Whose Job is it Anyway? A Study of Human-Robot Interaction in a Collaborative Task. *Human-Computer Interaction, 19:1/2*, 151-181.

Hoare, C.A.R. 1985. *Communicating Sequential Processes*. Prentice-Hall.

Krieg-Brückner, Bernd and Hui Shi. 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. International Conference GIScience 2006, Münster, Germany*. Berlin, Heidelberg: Springer, pp 234–250.

Moratz, Reinhard and Thora Tenbrink. 2006. Spatial reference in linguistic human-robot interaction: Iterative, empirically supported development of a model of projective relations. *Spatial Cognition and Computation,* Volume 6, Issue 1, pp. 63-106.

Pickering, Martin and Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 169-226.

Randell, D.A., Z. Cui, and A. Cohn. 1992. A spatial logic based on regions and connection. *Proceedings of the 3rd. International Conference on Knowledge Representation and Reasoning*, pp. 165-176, Morgan Kaufmann.

Ross, Robert J., Hui Shi, Tilman Vierhuff, Bernd Krieg-Brückner, and John Bateman. 2005. Towards Dialogue Based Shared Control of Navigating Robots. In Freksa, Christian, Markus Knauff, Bernd Krieg-Brückner, Bernhard Nebel, and Thomas Barkowsky (eds.), *Spatial Cognition IV: Reasoning, Action, Interaction*. International Conference Spatial Cognition 2004, Frauenchiemsee, Germany, October 2004, Proceedings. Berlin, Heidelberg: Springer, pp. 479-500.

Sakita, Tomoko I. 2006. Parallelism in conversation: Resonance, schematization, and extension from the perspective of dialogic syntax and cognitive linguistics. *Pragmatics & Cognition 14*:3, 467-500.

Schober, Michael F., and Susan E. Brennan. 2003. Processes of interactive spoken discourse: The role of the partner. In A.C. Graesser, M.A. Gernsbacher, and S.R. Goldman (eds.), *Handbook of Discourse Processes*, pp. 123-164. Mahwah, NJ: Lawrence Erlbaum Associates.

Shi Hui, Robert J. Ross, Thora Tenbrink, and John Bateman (subm.) Illocutionary Structure Modelling and the Analysis of Task-Oriented Dialogues.

Shi, Hui and Thora Tenbrink. 2009. Telling Rolland where to go: HRI dialogues on route navigation. In Kenny Coventry, Thora Tenbrink, and John Bateman (eds.), *Spatial Language and Dialogue.* Oxford University Press.

Sitter, S. and A. Stein. 1992. Modeling the Illocutionary Aspects of Information-Seeking Dialogues. *Information Processing and Management*, 28(2):165–180.

Steedman, M. 2000. *The syntactic process*. Cambridge, Massachusetts: MIT Press.

Taylor, Holly A. and Barbara Tversky. 1996. Perspective in spatial descriptions. *Journal of Memory and Language, 35,* 371-391.

Tenbrink, Thora. 2007. *Space, time, and the use of language: An investigation of relationships.* Berlin: Mouton de Gruyter.

Thrun, Sebastian. 2004. Toward a Framework for Human-Robot Interaction. *Human-Computer Interaction, Volume 19 (2004), Numbers 1 & 2*, pp. 9-24.

Vorwerg, Constanze and Thora Tenbrink. 2007. Discourse factors influencing spatial descriptions in English and German. In Thomas Barkowsky, Markus Knauff, Gérard Ligozat, and Dan Montello (eds.), *Spatial Cognition V: Reasoning, Action, Interaction.* Berlin: Springer, pp. 470–488.

Zoltan-Ford, Elizabeth. 1991. How to Get People to Say and Type what Computers Can Understand. *International Journal of Man-Machine Studies* 34: 527-547.