The Role of Users' Concepts of the Robot in Human-Robot Spatial Instruction

Kerstin Fischer

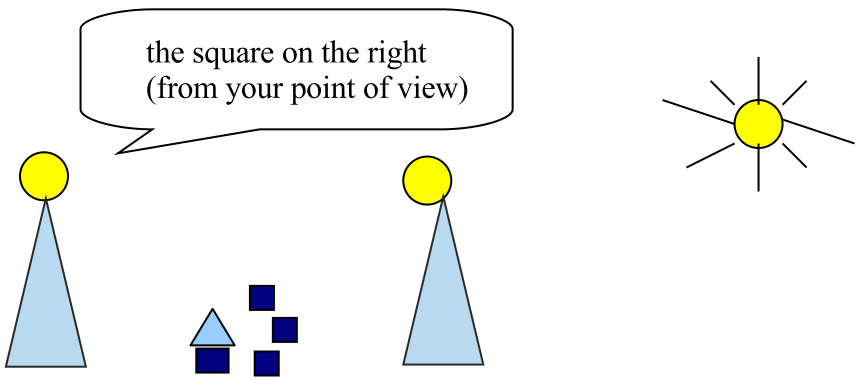
University of Bremen, FB10 – Sprach- und Literaturwissenschaften, 28334 Bremen
kerstinf@uni-bremen.de

Abstract. Spatial instructions are always delivered for a particular communication partner. In this paper I investigate the the role of users' concepts of their communication partner in human-robot interaction by analysing the spatial language choices speakers make in three comparable corpora with three different robots. I show that the users' concepts of their artificial communication partner is only mildly shaped by the appearance of the robot, and thus that users do not mindlessly use all clues they can get about their communication partner in order to formulate their spatial instructions. Instead, spatial instruction in human-robot interaction also depends on the users' models of the communication situation, as well as on external variables, such as gender.

Keywords: Spatial Language; Human-Robot Interaction; Linguistic Variation.

1 Introduction

The problem I address in this paper is the role of speakers' concepts of their communication partner in verbal spatial instruction. Speakers have a range of different ways of referring to spatial configurations at their disposal, and in a particular situation they have to decide for one variant over another. Among many other possibilities, speakers may refer to the spatial scenario in the following figure in different ways, for instance:



There are however multiple possibilities to refer to this, and to similar, spatial scenes, for instance:

the square on the right (from your point of view)	relative & partner-oriented
the square on the left (from my point of view)	relative & speaker-oriented
the square in front of the house	landmark-based & intrinsic
the square west of the house	landmark-based & extrinsic
the square in the middle	group-based
go straight, - a bit more to the right – stop	path-based

Table 1: Spatial Instruction Strategies

While these are not all possibilities for referring to spatial configurations, by far, they are the most relevant for our scenario discussed in the following. For a detailed overview of the spatial choices see Tenbrink (2005).

In this paper I address the conditions for this linguistic variation in spatial instruction focusing on the role of the communication partner. The question is thus, depending on whom speakers are talking to, which linguistic choices will they make?

Previous research on spatial language in dialogue has shown partner-specific effects particularly in the domain of perspective taking. Schober (1993) reports that if speakers had to imagine a communication partner, they exclusively relied on the partner's perspective, although Hermann and Grabowski (1994) have shown that this requires more cognitive effort for the speaker. If they had a real communication partner, they mostly used their own perspective. This is in line with results by von Stutterheim & Kohlmann (1998) who also found that in a situation in which the addressee could give feedback, speakers did not adapt their perspective. Schober (1998) attributes the difference to speakers' knowledge about the success of their instructions due to feedback. If their communication partners signal to them that understanding is successful, speakers may stick to the mode of instruction which needs the least collaborative effort (Clark & Wilkes-Gibbs, 1986). In case they do not know whether the communication partner understands them, they adjust to the partner's perspective in order to guarantee understanding in the absence of feedback. In Schober's (1995) study, however, speakers often take the perspective of their communication partner although there is the possibility of feedback. Moreover, after exchanging roles as instructor and instructee, speakers try to take the other's perspective as often as these have taken theirs. Finally, speakers' use of spatial instructions independent of particular viewpoints increases in these dialogues. Schober (1998) therefore argues that speakers are generally orienting towards the principle of ease in communication as a collaborative effort.

While these results point to partner-specific effects due to interaction, there are also results that show that the users' recipient design, their adapting their utterances to the expected needs of their communication partner (Sacks et al. 1974), influences the spatial choices made. Schegloff (1972), for instance, has shown that speakers may carry out extensive membership analyses in order to formulate a spatial description, that is, they try to determine the categories their addressee belongs to in order to infer the level of knowledge they can rely on in their spatial description. Similarly, Fischer and Bateman (2006) show that in human-robot communication speakers ask clarification questions particularly on these issues. Speakers' attribution of knowledge to their communication partner is furthermore apparent in Schober (2005) who shows that depending on the spatial capabilities of their addressees, speakers adapt to their communication partners. Moreover, Fischer (2005) shows that the choice of perspective in human-robot interaction depends crucially on what speakers believe the robot to be able to do. A linguistically skilful robot obviously creates a different, more complicated image than a robot that does not display any linguistic knowledge. The speakers' choice of perspective in these dialogues therefore depended essentially on the users' partner model.

What is still open, however, is the exact nature of the relationship between partner model, concept of the situation, external variables, and the spatial language chosen. My procedure here will be to analyse corpora of human-robot dialogues that differ only with respect to a single variable, namely the robot appearance. Human-robot dialogues, in contrast to natural dialogues among humans, have the advantage that the role of the addressee, the robot, can be investigated in much more detail than it can be done with humans since we can control the situations completely and vary only one variable at a time. Robots can moreover be manipulated far more easily than humans can be. Human-robot interaction, in contrast to human-to-human communication, allows therefore the identification of the parameters that speakers take into account when choosing their language for their particular communication partner (Fischer, 2003).

2. Method and Data

In the framework of the I1-project in the SFB/TR8 in Bremen, we elicit human-robot interaction data in scenarios that differ only with respect to single parameters. On the basis of the data elicited, we determine the influence of each parameter on the users' linguistic choices, and identify correlations in users' choices in order to predict behaviour on the basis of non-intrusive user models. The corpora used in this study are dialogues in which only the robot appearance is varied, and all other factors are kept constant. In the methodology employed here (Fischer, 2003), the comparability and control of the situation is achieved by keeping the robot behaviour constant. That is, the robot's behavioural (and, in other corpora, verbal) output is based on a fixed schema which is the same for all dialogues and across corpora. In this way, not only are all speakers confronted with exactly the same situation. The methodology also allows us to investigate speakers' concepts of their (artificial) communication partner. Because of the fixed schema that is causally unrelated to the speakers' actual linguistic behaviour, the speakers' sense-making efforts cannot be attributed to particular features or misbehaviours of a particular robot, but they have to be understood as arising from the speakers' own cognitive models of the situation, including their communication partner.

For example, if the speaker after a miscommunication (which is of course frequent if the robot does not really react to the speakers' utterances) uses a different descriptive term for the goal object, she displays that she holds the problem to be due to the robot's limited lexicon. If she switches to another reference system, she displays that she believes the reference system to be possibly inappropriate in this situation. That is, the original utterance and its reformulation serve as solutions to the same communicative task, and the second choice is based on an analysis what may have gone wrong in the previous turn, depending on the speaker's concept of the communication partner. Admittedly, a single instance of a particular linguistic choice may be attributed to many other factors as well; it is the systematicity and the repeatedness both for the same speaker and between speakers that allow the conclusion that we are observing the users' strategies depending on their concepts of their communication partner. To conclude, the design allows the identification of the explanatory models speakers build up to make sense out of the human-robot situation (Fischer, 2003).

Fig. 1: Aibo

Finally, the frequent impression of miscommunication encourages speakers to employ more than one linguistic strategy. If speakers are immediately successful, they have been found to stick to this strategy (Moratz et al., 2001, Fischer, 2006). This would however yield very uninteresting data to us. Instead, the speakers' reformulations are particularly revealing concerning the participants' concepts of the communication partner and their understanding of the situation.

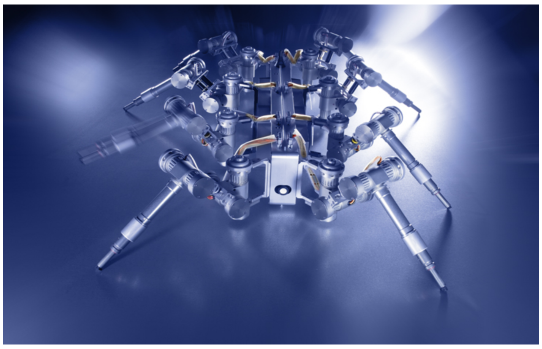


Fig. 2: Scorpion

The dialogues discussed here differ only with respect to the robot addressed. The first robot used was Sony's Aibo, a commercially available robot which looks like a little dog (see Fig 1). The second robot has been built by Frank Kirchner and colleagues at the University of Bremen (e.g. Spenneberg and Kirchner, 2002) and looks like a metal spider (see Figure 2). The third robot used is another commercially available experimental robotic platform, the Pioneer (see Fig 3).

Procedure The task users had to fulfil was to verbally instruct the robot to move to particular goal objects, pointed at by the experimenter. The objects, as well as the robot, were placed on the floor in front of the participants in various spatial configurations, changed during the experiment. Pointing was used in order to avoid prompting the participant with particular spatial descriptions. Most tasks involved a single goal object, one task involved a sequence of several objects to which the robot had to go. After two tasks, which took up about two thirds of the time per dialogue, the experimenter told each speaker that she was allowed to refer to the objects directly. This prompt was meant to investigate the speakers' willingness to change their instructions, given their previous experience with the robot. The robot was steered by a student employee behind a screen according to a fixed schema of robot behaviours (Fraser & Gilbert, 1991). After the recording, speakers filled out a questionnaire in which they were asked whether they had believed that they were talking to a real robot, which all of them acknowledged.

Fig. 3: Pioneer

Participants Participants were 66 German students from the University of Bremen of whom 12 interacted with Aibo, 21 with Scorpion, and 33 with Pioneer. Each participant interacted with just one robot, going through all task, which are defined by the various spatial configurations that we tested. Thus, all received prompts after the second task, and all filled out questionnaires at the end of the session. The instruction was simply to make the robot move to the objects pointed at by the experimenter. The prompt was 'oh, by the way, you may refer to the objects directly'.

The data elicitation conditions can thus be summarised as follows:

- 66 German human-robot dialogues
  - 12 dialogues with Aibo, a dog-like pet robot
  - 21 dialogues with Scorpion, a huge metal insect (e.g. Spenneberg and Kirchner, 2002)
  - 33 dialogues with Pioneer, a car- or box-like robot
- Task: to instruct the robot verbally to move to particular goal objects
- Robot behaviour: according to fixed schema of behaviours, independent of the speakers' utterances ('Wizard-of-Oz' scenario)
- after two tasks (after about 10 minutes of interaction), speakers were prompted that they can use object-based descriptions

The statistical analysis ANOVA was carried out using the Statistica software package.

Data Coding The users' linguistic behaviour in the first two tasks, before the prompt, were coded in the following way:

- different types of **spatial instruction**, in particular, path- versus object-based instructions; that is, the speakers' linguistic behaviour in the first two tasks, before the prompt, was analysed with respect to consistent use of path-based instructions, such as '*move left*', consistent use of goal-based instructions in which the participants name an object, such as '*go to the left bowl*', or mixed usage;
- different degrees of **emotional expression** as evidenced by the use of interjections, partner-oriented signals and displays of interpersonal relationship, like *my friend*;
- different types and amounts of **feedback**, for instance, whether the speaker used feedback signals, like *okay*, explicit evaluations, like *good job*, and the use of character traits, e.g. *good boy*;
- different assumptions of competence; here, we coded the amount of **structuring cues** (see also Fischer & Bateman, 2006), distinguishing between implicit, for instance, *now*, and explicit, for instance, *first of all, the next step,* structuring cues as well as task overviews. Moreover, we coded different types of **intensifiers**, in particular, colloquial, for example, *slightly,* versus technical, for instance, *30 degrees*, and metaphorical, for example, *sharp*, intensification. Finally, we investigated the abstractness of **object descriptions**, for example, whether users

employed abstract terms, such as *object* or *obstacle*, or whether they used basic-level terminology, for instance, *bowl*;

- different displays of relationship between the communication partners, for instance, with respect to **anthro- and zoomorphisation**, **politeness**, e.g. the use of *please* and *thank you*, **reference to the robot**, for instance as *robot* and *it* or as *he*, and the **sentence mood** chosen, in particular, whether no verb is used or whether the verb is in the infinitive and thus no assertion about the relationship is made, or whether the imperative is chosen or other linguistic forms, for example, the declarative or complex modal constructions;
- different intonation contours, indicating either different degrees of certainty and assertiveness or different degrees of interactivity. That is, rising **intonation contours** have been suggested to indicate lack of certainty (Lakoff, 1972), or to function as a device to increase interactivity and to encourage the partner to contribute to the conversation (Fishman, 1983). Thus, spatial instructions were coded for whether prosodic delivery was overwhelmingly with rising, falling, or level intonation contour, or whether it was mixed;
- different displays of the concept of the human-robot interaction situation as evidenced by the **dialogue beginning**, in particular, whether the speakers begin the interaction with a greeting or whether they immediately proceed to an instruction.

3. Hypotheses

There is ubiquitous research showing that speakers orient at their communication partners' appearance when formulating their utterances. Not only Schegloff's (1972) findings on membership analysis make us expect that speakers will use every clue possible to create models of their artificial communication partners that help them make their linguistic choices; also Roche (1989), for instance, has shown that appearance is a crucial factor in intercultural communication. Moreover, results by Nass and colleagues (Nass and Moon, 2000; Nass and Brave, 2005) show that users transfer knowledge from the source domain of human social interaction to human-computer interaction. All these findings make us expect that the spatial language directed at robots that differ in appearance will differ considerably. The hypothesis investigated here is thus that the users' spatial instruction strategies should crucially depend on their concepts of their artificial communication partner, which in turn should be influenced by the appearance of the robot. The analysis should thus concentrate not only on the spatial language features themselves, but also on evidence of the users' concepts of their communication partner as apparent from other linguistic choices. The prediction, based on previous findings, is thus that the linguistic features outlined will be distributed differently in the different corpora.

Besides the prediction that the users' linguistic behaviour will differ significantly in general, we may furthermore try to identify in more detail what the speakers' concepts are determined by. In particular, we may ask whether the users' attributing properties to robots takes place on the basis of analogies from the source domains of the robots, i.e. the animal kingdom and, for example, toy cars. If this is the case, it can be predicted that with respect to the features analysed, users will ascribe more competence to a more technically appearing system, here the Pioneer, compared to robots whose source domains are from the animal kingdom. Correspondingly, spatial instruction strategies, structuring cues, intensifiers, and object-naming strategies may be distributed differently for the three robots, such that Pioneer is trusted with a higher competence and Aibo and Scorpion with a lower degree of competence.

In contrast, those features pointing to stronger interpersonal and emotional relationships between speaker and robot should be higher for Aibo and possibly the Scorpion than for Pioneer. Thus, if speakers make use of properties of the source domains, one may expect more zoomorphisation for Aibo and Scorpion than for Pioneer, more instances of personifying references to the robot for Aibo and Scorpion than for Pioneer, more feedback, more emotional expression, more interactivity with Aibo and Scorpion, as evidenced by the prosodic delivery of the instruction, and more interpersonal information in the sentence mood and in the dialogue beginning.

4. Results

4.1. General Results

The first result is that without extra information, speakers consistently used only partner-centred instructions; that is, they exclusively took the robot's perspective, sometimes even getting up from their chair to turn into the direction which the robot was facing in order to identify the reference frame from the robot's point of view. Exceptions are the few goal object-based descriptions, in which the reference frame was often group-based. Thus, the spatial instruction style found was in 42.4% of all cases exclusively, and in 77.2% mostly, path-based. After the prompt, that speakers may also refer to the objects directly, however, 79% switched to object-based descriptions, most of which were group-based, which shows that speakers generally prefr this instruction strategy, yet did not initially believe it to be useful in the human-robot situation. The spectrum of spatial strategies taken throughout the dialogues comprises:

path-based	*go right*
goal-based	*the bowl on the right* (meaning the robot's right)
group-based	*to the bowl in the middle* (meaning in the middle of a group of objects)

4.2. The Role of the Robot

The analysis of variance shows that there are significant differences with respect to **spatial instruction** strategies between the corpora ($F=11.08$; $df=2$; $p < 0.002$). In particular, post hoc analyses show that whereas spatial instruction for Aibo is mainly path-based, the numbers of object-based instructions is considerably higher for Pioneer and Scorpion than for Aibo (see Figure 4).

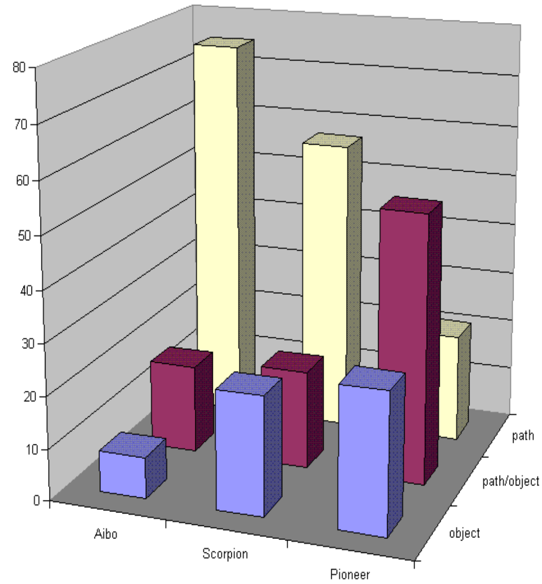


Fig. 4: Percentages of Spatial Instruction Strategies Used for the three Robots

The t-test (one-tailed) reveals that users employed significantly more higher-level (object-based) spatial instructions for Pioneer than for either Aibo or Scorpion. That is, Aibo and Scorpion are addressed with similar, very basic constructions, whereas Pioneer is addressed in a more elaborate way. Details are shown in the Table 2:

robot	Aibo	Scorpion	Pioneer
Aibo	-	n.s.	$p<0.003$
Scorpion		-	$p<0.07$
Pioneer			-

Table 2: Differences in Spatial Instruction

Furthermore, the choice of **intensifiers** differed for the different robots (F=3.877; df=3; $p < 0.03$). The four different types of intensifiers identified in the corpora were **technical**, such as *30 degrees to the right*, **colloquial**, such as *slightly*, **metaphorical**, such as *sharp*, or **none** at all (see Fig. 5).

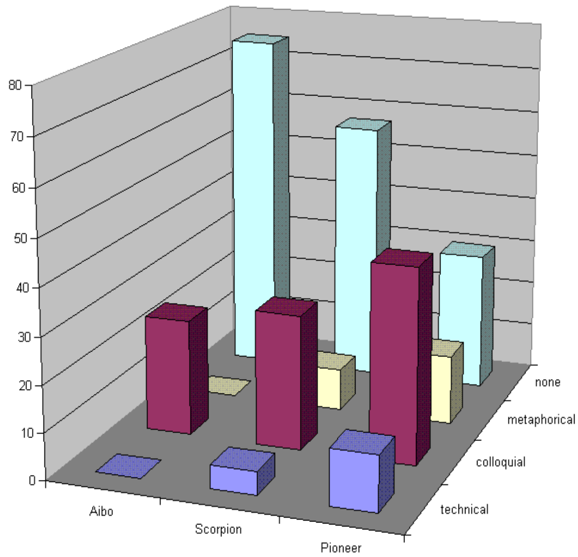


Fig. 5: Percentages of Intensifier Types Chosen for the three Robots

Post-hoc analysis by means of the t-test shows that users employed significantly more intensification for Pioneer than for either Aibo or Scorpion. That is, again Aibo and Scorpion are addressed in similar, more restricted ways, whereas Pioneer is addressed in a more elaborate way. Details are shown in the Table 3:

robot	Aibo	Scorpion	Pioneer
Aibo	-	n.s.	p=0.09
Scorpion		-	n.s.
Pioneer			-

Table 3: Differences in Intensification

Furthermore, the **prosodic delivery** was revealed to be significantly (F=15.1488; df=1; $p < 0.01$) different for the three robots. In particular, post-hoc analysis shows that speakers delivered their spatial instructions for Aibo with falling intonation contours, while for Pioneer they employed rising intonation contours, see Table 4.

Scorpion is somewhat in the middle, since it also differs significantly from the Pioneer, but with a statistical tendency for a difference with Aibo as well.

robot	Aibo	Scorpion	Pioneer
Aibo	-	p=0.11	p<0.0002
Scorpion		-	p<0.003
Pioneer			-

Table 4: Differences in Intonation Contours

Unfortunately, this result is difficult to interpret due to the multifunctionality of intonation in discourse. Thus, a rising intonation contour may express uncertainty, in this case supporting the hypothesis that speakers orient more to the 'devicehood' of the more mechanically appearing robots Scorpion and Pioneer, suggesting that users are uncertain about communicating with such technical devices. However, rising intonation also has the function of encouraging interactive negotiation, and thus it could be argued that speakers are more interactive with Pioneer and Scorpion. Finally, also the source domain hypothesis cannot be ruled out completely since with dogs we are used to use falling intonation contours.

The most striking result however is that contrary to expectations, regarding all other variables investigated none turned out to be significantly different. In particular, none of the variables **emotional expression, feedback, structuring cues, object abstractness, zoomorphisation, politeness, reference to the robot, grammatical mood** or **greeting** yielded statistical differences between the corpora. This is particularly surprising for the amount of zoomorphisation, which should be higher for zoomorph robots like Aibo and Scorpion, yet speakers turned out to use similar conceptualisations for the Pioneer, for example:

(1) P110: den mag er besonders gern nicht? -- (laughter), den hat er zum fressen gern, (laughter) (breathing) [*this one he likes best, doesn't he? (laughter), he could eat it alive (laughter)*] [ol4,box2]

Thus, many factors related to the attribution of competence and emotionality regarding interactional management and interpersonal relationship were not influenced by the appearance of the robots.

4.3. The Role of External Variables: Gender

Investigating the corpora further for other relevant factors reveals that classical sociolinguistic variables are of some influence. The speaker variable gender turned out to be the most relevant regarding some of the variables suspected to be involved in different displays of relationship between human and robot. Thus, reference to the robot (F=5.3216; df=1; p < 0.03), politeness (F=5.7493; df=1; p < 0.03) and emotionality (F=3.3604; df=3; p < 0.02) were significantly influenced by the speakers' sex, such that men referred to the robot less often in general, and less often in an zoomorphised way, i.e. by using he or boy, for instance. Furthermore, men were found to use fewer interjections and contact signals and fewer politeness items, such as *thank you* or *please*.

Moreover, a linguistic feature showing different degrees of competence ascribed to the robot, the choice of object-naming strategies, was also found to be essentially gender-specific (F=5.3722; df=2; p < 0.01). In particular, women used more concrete object descriptions, such as *bowl, glass, pot*, instead of abstract terms, such as *object* or *obstacle*.

4.4. The Role of Participants' Concepts of the Communication Situation

The users' choice of dialogue beginning, which is suspected to indicate the speakers' concept of what the particular human-robot interaction situation consists in (see Fischer 2006ab), proved to be influential for the choice of structuring cues (F=3.12678; df=3; p < 0.04). That is, speakers who greeted the robot in some way or other were more inclined to use structuring cues, such as *next, then*, or even to give descriptions of the spatial scene, e.g. *vor Dir sind drei blaue Schälchen zu sehen – fahr mal zu dem mittleren. -- blauen Schälchen* [*in front of you are three small blue bowls. why don't you drive to the middle – blue bowl.*], than those who immediately start the dialogue with a spatial instruction. Thus, the use of a greeting at the dialogue beginning correlates with attributions of higher linguistic, perceptual, and cognitive competence.

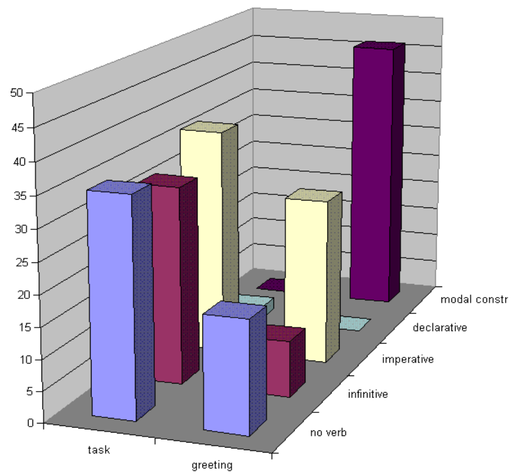


Fig. 6: Percentages of Sentence Mood Used by Speakers who Begin the Dialogues with a Greeting versus with a Task-oriented Instruction

Moreover, the users' dialogue beginnings also correlate with their choice of sentence mood ($F=6.426$; $df=2$; $p < 0.003$). That is, speakers who greet the system are more likely to use a verb in their spatial instruction (instead of using only spatial terms, for instance, *right*, *straight*) and to employ grammatical mood that expresses an interpersonal relationship between speaker and hearer, such as the imperative, the declarative mood or even complex modal constructions, such as *why don't you* in the example above. In contrast, speakers who begin their dialogues with a spatial instruction are more likely to use the infinitive, the imperative, or no verb at all (see Figure 6).

5. Discussion

Human-robot interaction constitutes for most speakers a very unusual communication situation, and thus it could have been expected that speakers use all information available to them from the robot to design their utterances for their unfamiliar communication partner. Accordingly, Nass and Brave (2005) have found that speakers employ the cues given by the voice of human-computer interfaces to infer properties of their communication partner. They propose that people mindlessly transfer knowledge about human social actors to human-computer interfaces (see also Nass and Moon 2000). Starting from this hypothesis I expected that users in our human-robot dialogues would transfer many properties of the robot source domain to the verbal interaction. However, this was the case only to a very limited degree. Indeed the spatial instruction strategies were adapted to the particular robot such that the pet robot Aibo and the Scorpion were addressed in a very low-level style of spatial instruction and with no or only colloquial intensification. Moreover, Aibo and Scorpion were found found to be addressed with falling intonation, in contrast to the Pioneer, which may correspond to the fact that one orders dogs to go somewhere, rather than to ask them; however, it may also be related to uncertainty in the communication with artificial communication partners, Pioneer being more 'robotic' in its appearance than Aibo.

However, if the source domain was what mattered the most, a Scorpion would probably not be hypothesised to be as intelligent as a dog. Indeed, Scorpion was addressed with instructions at the same level as Aibo and with similar amounts of technical and metaphorical means of intensification.

Moreover, due to the fact that Aibo is a pet robot, emotional expression and zoomorphisation would have been expectable to a higher degree if users really mindlessly transferred properties of human-dog communication to human-Aibo interaction. However, the most important influencing factor here was gender, not the difference between Aibo versus Scorpion or Pioneer.

Finally, if the source domains mattered that much, also other linguistic properties should have been influenced by the appearance of the robots. Instead, only three linguistic properties, spatial instruction, intonation contour and intensification, were found to differ significantly between the robots, whereas most features relevant in the interpersonal relationship and the attribution of competence to the robot were not affected.

The explanation for these findings, I propose, can be found in the fact that in the dialogues investigated here, speakers are talking to an artificial communication partner more than to anything else. The hypothesis would thus be that instead of relying on cues from the source domains *doghood, scorpionhood* or *technical devicehood* for the three robots, I want to suggest that the speakers in fact reacted to different degrees of artificialness of the robots, which in turn influences the definition of the human-robot situation in general.

Moreover, in the English data recorded in the project under the same conditions, none of the linguistic features investigated differed significantly for the three robots (Fischer, forthcoming), not even instructional strategy and intensification. This may be partly due to the fact that there we used the same speakers for all three robots, although recording took place at least three months apart. However, the lack of differences fits the overall picture very well, indicating a rather weak role of the appearance of the robot. However, also in the English data, the linguistic choices users made correlated significantly with the dialogue beginnings, as it was found for the data analysed here with respect to structuring cues and grammatical mood. Thus, structuring cues and sentence mood, variables that had been suspected to be related to ascribed competence of the communication partner and the interpersonal relationship respectively, have been found to be determined by the way speakers conceptualise the communication situation as it is apparent from their way of opening the dialogue. This is in line with previous findings in human-robot and human-computer communication, which show that the speakers' own perception of the communication situation determines the linguistic properties of their utterances, e.g. the use of clarification questions (Fischer and Bateman 2006), the reaction to situations of misunderstanding (Fischer 2006a, Fischer and Wilde 2005), the prosodic delivery of utterances (Fischer 2006a), as well as cooperative and uncooperative conversational strategies (Fischer 2006a). With respect to these linguistic strategies it was found that they can be significantly related to the users' conceptual models of the communication situation as either managing a technical tool or as pretending to enter a conversation with the artificial communication partner (see also Fischer 2006b).

Considering the relationship between partner model, concept of the situation, external variables, and the spatial language chosen, we can conclude that all three play a role in the formulation of spatial utterances. While the concepts evoked by the particular robots interact with concepts of the human-robot communication situation in general, and in particular with the concept of the robot as human-like or tool-like, in creating a model of the partner with respect to its capabilities, external factors, here gender, also play a decisive role with respect to interpersonal factors of human-robot interaction.

References

Clark, Herbert H. & Wilkes-Gibbs, Deanna (1986). Referring as a collaborative process. *Cognition* 22, 1-39.

Fischer, Kerstin (2003). Linguistic Methods for Investigating Concepts in Use. In Stolz, Thomas; Kolbe, Katja (eds.), Methodologie in der Linguistik. Frankfurt a.M.: Peter Lang.

Fischer, Kerstin (2005): Discourse Conditions for Spatial Perspective Taking. *Proceedings of the Workshop on Spatial Language and Dialogue*, WoSLaD, Delmenhorst, October 2005.

Fischer, Kerstin (2006a). *What Computer Talk Is and Isn't. Human-Computer Conversation as Intercultural Communication.* Saarbrücken: AQ-Verlag.

Fischer, Kerstin (2006b): The Role of Users' Preconceptions in Talking to Computers and Robots. *Proceedings of the Workshop on How People Talk to Computers, Robots, and other Artificial Communication Partners*, Delmenhorst, April 21-23, 2006.

Fischer, Kerstin (forthcoming): Talking for an Addressee. Alignment, Recipient Design and Interaction in So-called 'Simplified Registers'. Habilitation Thesis, University of Bremen.

Fischer, Kerstin & Wilde, Ruth (2005). Methoden zur Analyse interaktiver Bedeutungskonstitution. In Ueckmann, Natascha (ed.), Von der Wirklichkeit zur Wissenschaft: Aktuelle Forschungsmethoden in den Sprach-, Literatur- und Kulturwissenschaften. Hamburg: LIT.

Fischer, Kerstin & Bateman, John A. (2006): Keeping the initiative: An empirically motivated approach to predicting user-initiated dialogue contributions in HCI. *Proceedings of EACL 2006*, Trento, Italy.

Fishman, Pamela (1978): Interaction: The work women do. *Social Problems* 25: 397-406.

Fraser, N. and Gilbert, G.N. (1991). Simulating Speech Systems. *Computer Speech and Language* 5, 81-99.

Hermann, Theo and Grabowski, Joachim (1994). Sprechen. Psychologie der Sprachproduktion. Heidelberg: Spektrum.

Lakoff, Robin (1975): *Language and Woman's Place*. Oxford University Press.

Moratz, Reinhard; Fischer, Kerstin and Tenbrink, Thora (2001). Cognitive Modelling of Spatial Reference for Human-Robot Interaction. In *International Journal on Artificial Intelligence Tools* 10,4, 589-611.

Nass, Clifford & Brave, Scott (2005): *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship.* Cambridge, MA: MIT Press.

Nass, Clifford & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81-103.

Roche, Jörg (1989): Xenolekte. Struktur und Variation im Deutsch gegenüber Ausländern. Berlin/New York, de Gruyter.

Sacks, Harvey, Schegloff, Emanuel A. and Jefferson, Gail (1974): A Simplest Systematics for the Organization of Turn-Taking for Conversation, *Language* 50: 696-735.

Schegloff, Emanuel A. (1972). Notes on a Conversational Practice: Formulating Place. In D. N. Sudnow (ed.), *Studies in Social Interaction*. New York: MacMillan, The Free Press, 75-119.

Schober, Michael F. (1993). Spatial Perspective-taking in Conversation. *Cognition* 47, 1-24.

Schober, Michael F. (1995). Speakers, Addressees, and Frames of Reference: Whose Effort Is Minimized in Conversations about Locations? *Discourse Processes* 20, 219-247.

Schober, Michael F. (1998). How Addressees Affect Spatial Perspective Choice in Dialogue. In Olivier, Patrick & Gapp, Klaus-Peter (eds.), *Representation and Processing of Spatial Expressions*. London: Erlbaum, (pp. 231-245).

Schober, Michael F. (2005): Spatial Dialogue between Speakers with Mismatched Abilities. Proceedings of the Workshop on Spatial Language and Dialogue, WoSLaD, Delmenhorst, October 2005.

Spenneberg, Dirk & Kirchner, Frank (2002): "Scorpion: A Biomimetic Walking Robot", In: Robotik 2002, Vol. 1679, VDI-Bericht, pp. 677-682.

Tenbrink, Thora (2005): Localising objects and events: Discoursal applicability conditions for spatiotemporal expressions in English and German. PhD Thesis, University of Bremen.

von Stutterheim, Christiane & Kohlmann, Ute (1998). Selective Hearer-Adaptation. *Linguistics* 36, 3, 517-549.