2008 IEEE International Conference on
Robotics and Automation
Pasadena, CA, USA, May 19-23, 2008

"Try something else!" - When users change their discursive behavior in human-robot interaction

Manja Lohse, Katharina J. Rohlfing, Britta Wrede, and Gerhard Sagerer, *Member, IEEE*

Abstract— **This paper investigates the influence of feedback provided by an autonomous robot (BIRON) on users' discursive behavior. A user study is described during which users show objects to the robot. The results of the experiment indicate, that the robot's verbal feedback utterances cause the humans to adapt their own way of speaking. The changes in users' verbal behavior are due to their beliefs about the robots knowledge and abilities. In this paper they are identified and grouped. Moreover, the data implies variations in user behavior regarding gestures. Unlike speech, the robot was not able to give feedback with gestures. Due to the lack of feedback, users did not seem to have a consistent mental representation of the robot's abilities to recognize gestures. As a result, changes between different gestures are interpreted to be unconscious variations accompanying speech.**

I. INTRODUCTION

WHETHER interacting with a colleague from another department, a child with distinct cognitive and linguistic skills or a foreigner with different cultural background, humans try to adapt to their communication partners. In everyday interaction this process seems to happen automatically. Humans send certain clues and give verbal and nonverbal feedback. With their help the interaction partners form mental representations of each other. They build up beliefs about the others abilities and knowledge. By doing so, humans are able to adapt to others, which increases the possibility that the interaction is successful. The same is true when the interaction partner is a robot. Especially with the development of so-called social robots it is increasingly important to know more about peoples' beliefs about the robot in order to design a successful interaction. Users' beliefs can be studied by analyzing their behavior in a certain interaction situation. Knowing how users behave, moreover, helps to design dialogs. In this paper, a study with the service robot BIRON ((BIelefeld RObot companioN; see Fig. 1) is presented which aims at shedding some light at the user behavior in the situation of teaching objects to a robot.

Manuscript received September 14, 2007. K. Rohlfing's work was supported in part by the Volkswagen Foundation.

M. Lohse, K. J. Rohlfing, and B. Wrede are with the Applied Computer Science Group, Bielefeld University, Bielefeld, Germany (phone +49 521 1062953; fax: +49 521 1062992; e-mail: mlohse; rohlfing; bwrede@techfak.uni-bielefeld.de).

G. Sagerer is head of the Applied Computer Science Group, Bielefeld Uniersity, Bielefeld, Germany (e-mail: sagerer@techfak.uni-bielefeld.de)

Fig. 1 Person interacting with BIRON

II. RELATED WORK

Since human-robot interaction is a rather young research field based on various disciplines, many approaches and scientific findings from other areas have to be taken into account. This is also true for related work on human discursive behavior which has been in the focus of human-computer interaction. [1] report research on linguistic adaptation during spoken and multimodal human-computer interaction in situations when errors occur. Their work focuses on modalities and intonation. The researchers conclude that users adapt to the system in three different ways: increasing linguistic contrast (alternation of input mode and lexical content); increasing hyperarticulation; suppression of linguistic variability (amplitude and frequency) when hyperarticulating. This work focuses on linguistic phenomena. [2, 3] also concentrate on error recovery in spoken dialog systems. In contrast to [1], the ten error recovery strategies the authors propose include non-linguistic phenomena. Their research focuses on the feedback of the system, whereas the work presented here concentrates on users' utterances and gestures in reaction to the robot.

Feedback as such plays a major role in HRI. According to [4] the term feedback describes "linguistic mechanisms, which enable the participants in a communication process to unobtrusively exchange information about four basic communication functions: contact, perception, understanding and attitudinal reactions". Thus, it is the basis for grounding. [5] define the common ground on the level of speech and context. To establish a common ground the interaction partners have to go through a knowledge estimation process. Based on this idea, users usually adapt to the robots motivated by the systems linguistic output and behavior [6]. Thus, it can be expected that humans try various communication behaviors during the interaction. These behaviors point to their models of and attitudes towards the system [5, 7, 8]. In order to achieve an enjoyable interaction they have to be shaped in a way that allows the user to build a functional model of the system [9]. This model should also help to improve turn-taking which is an important factor in human-robot interaction as well as in human-human interaction [10, 11, 12].

III. EXPERIMENT

A. System and Scenario

The robot used in the experiment reported here is called BIRON. It is developed as a service robot for the household. Thus, the goal is to implement a multimodal situated system that can learn the spatial environment as well as the names and visual appearance of objects. Therefore, BIRON cannot only understand spoken utterances but also co-verbal deictic references to objects in the scene. Moreover, the robot can carry out mixed-initiative dialogs. For a detailed description of the system see [13].

The development of BIRON is embedded in the so-called home tour scenario. In this scenario the robot is delivered to an inexperienced user's home where it has to be introduced to its new environment. Thus, the user shows the robot around and teaches it important places and objects. For this scenario, which addresses human-robot interaction with novice users, intuitive robot operation is indispensable. That's why live trials with inexperienced subjects, like the one presented here, are helpful and necessary.

B. Method, Setup and Users' Task

To proof the hypothesis that users change their discursive behavior (speech and gesture) depending on the feedback of the robot, a study with 15 German native speakers was conducted.

Before interacting with the robot, subjects had to fill in a questionnaire providing information about their age, gender, experience with robots and computers. Moreover, they were asked to check which of the given robots (Aibo, Kismet, mars explorer, Asimo, soccer robot, Lego Mindstorms, Roomba, R2D2, BIRON, service robot for the home) they knew. The attitude towards robots was further studied by asking how much people liked the idea of a household robot that finds its way around an apartment and accomplishes certain tasks in everyday-life just like a butler would do. Participants also rated whether they would want to have such a robot at home themselves. The last questionnaire item was a rating of the importance of different abilities the robot should have (understand speech, recognize gestures etc.).

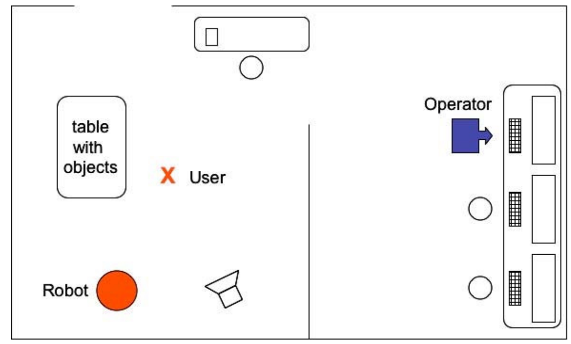

Fig. 2 Setup of the experiment

After answering this questionnaire users received written instructions about the task they had to perform. Inspired by the home tour scenario, people had to show about five objects to the robot and to teach it their names. This task is closely related to the home tour scenario since learning about the environment is a mandatory ability a domestic robot has to have. Subjects could chose between different objects: a bottle, a cup, a remote control, a fork, a spoon, a pencil, a book, and a watering can. All objects are suitable for a household scenario.

Participants were told that the robot needed some more time for processing than humans do. Apart from that, they were instructed to talk only to the robot during the interaction and to try to solve upcoming problems with BIRON.

During the experiments the robot operated in its fully autonomous mode, which was necessary to produce realistic communication sequences including problems caused by the complex interaction of the diverse perceptual system components. Only speech recognition was simulated by manual text input to improve the recognition rate and to speed up the robots reactions. The operator typed in all user utterances directed to the robot. Peculiarities that most speakers have (e.g. cutting of words) were ignored. Altogether, all utterances were typed in like they would occur in written language. Moreover, expressions like "hmmm" were not typed in, since they don't have any meaning to the robot.

The experiments took place in a robotic lab. Fig. 2 gives an impression of what the setup looked like. Since a wall separated the room, participants didn't notice the operator during the interaction with the system.

Next to general expressions like "Hello", the robot used a limited set of feedback utterances regarding the task: That's interesting. I really like it.; Yes please?; I just understood you partially. Can you please repeat?; Pardon.; Sorry, I'm still young and can't do this.; Sorry, I can't search for objects right now.; Sorry, I don't know. This restricted set of answers, along with the concrete task, allowed for a comparison of user behaviors. The interaction was videotaped. Afterwards, subjects filled in another questionnaire judging the robot and its abilities.

C. *Analysis of Data*

The videos were transcribed with the help of ELAN. Transcriptions were made of

- Utterances of the human
- Utterances of the robot
- Overlap between utterances (interruptions)
- Gestures of the human
- Objects shown to the robot

The transcript was then analyzed and behaviors (speech and gesture) of the human were identified and grouped. For speech this was done with the help of a linguistic analysis. Analyzing the videos, only units of speech and gesture that convey meaning concerning the task were taken into consideration. Thus, utterances like "mhm" or the scratching of the chin were not interpreted as a conscious behavior conducted to show an object to the robot. With the help of the video analysis eight task-related discursive behaviors were identified:

1. naming object (whole sentence)
   example: "This is a cup."
2. naming object (one word, very short utterance)
   example: "Cup"
3. describing the object
   example: "The cup is blue and has a handle."
4. asking for feedback regarding the object
   example: "BIRON, do you know what this is?"
5. asking for BIRON's general abilities and knowledge
   example: "BIRON, what can you do at all?"
6. asking for BIRON's ability to listen / speak
   example: "BIRON, can you hear me?"
7. asking for BIRON's ability to see
   example: "Can you see the object?"
8. demanding attention for the user / object / task
   example: "BIRON, look at me."

A change was annotated as soon as the participant switched between behaviors.

Next to typical verbal behaviors the data also implies some patterns concerning task-related gestures the subjects used. We propose nine groups of gestures:

1. Presenting the object (Fig. 3A)
2. Moving the object once (up, down, to another position, rotate)
3. Moving the object continuously (back and forth, up and down, to different positions, rotate back and forth)
4. Moving the object closer to the robot
5. Manipulating the object (open the book / bottle)
6. Looking at the object
7. Pointing at the object (Fig. 3B)
8. Imitating actions that can be performed with the object (drinking, eating, reading etc.) (Fig. 3C)
9. Holding the object (Fig. 3D)

Fig. 3 A) Presenting the object; B) Pointing at the object; C) Imitating action; D) Holding the object

IV. RESULTS

In the following section the results of the experiments are presented. It starts with the evaluation of the questionnaires. Thereafter, the videos are analyzed with the help of the verbal and gesture behaviors mentioned above.

A. *Results of the questionnaires*

All 15 participants were students. Their age ranged between 22 and 37 with an average of 24.7 years. Even though the majority had some experience working with computers (mean 3.3 on a scale of 1 to 5) only one person indicated that she had some minor experience interacting with robots. Apart from this self-assessment it was asked which of the following robots the subjects knew. Fig. 4 shows the results. Out of the 10 robots given, participants knew an average of 3.3.

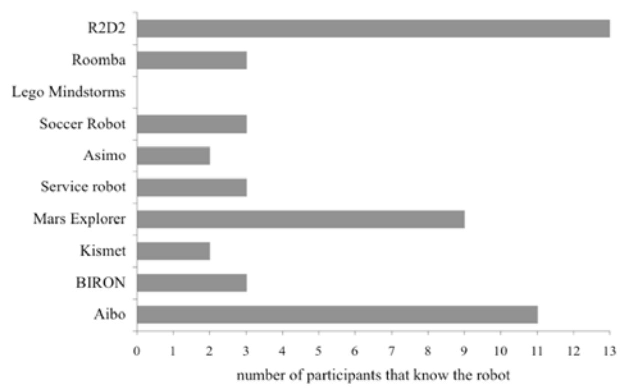

Fig. 4 Robots known by the participants

Subjects were inexperienced users in the sense that they had no former experience interacting with robots. Nevertheless, they had gained some knowledge about robots, obviously from the media. Asked before the

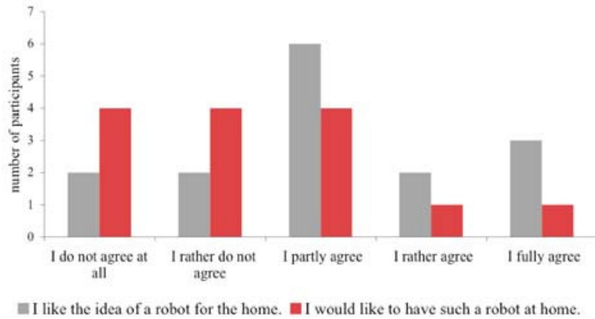

Fig. 5 Ratings of the questions: I like the idea of a robot for the home; I would like to have such a robot at home

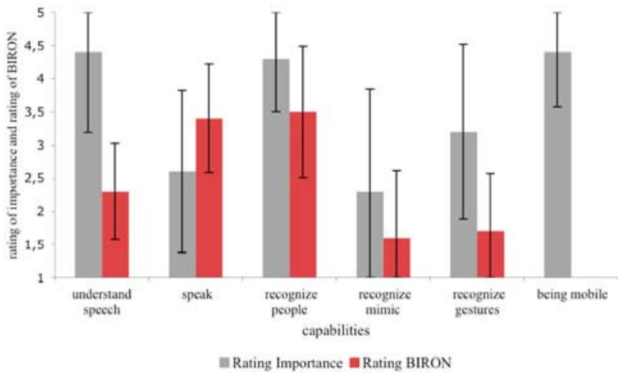

Fig. 6 Importance of capabilities and rating of BIRON's capabilities

interaction with BIRON, the majority of the participants were more or less open to the idea of robots in the home in general (m=3.1 on a scale of 1 to 5) (Fig. 5). Nevertheless, they did not feel very positive about having such a robot at home themselves (m=2.2 on a scale of 1 to 5).

A robot for the domestic environment needs certain abilities in order to fulfill its tasks. Thus, participants were asked to rate on a 5-point scale how important they thought the following abilities were: understand speech, speak, detect persons, recognize mimic, recognize gestures and being mobile. Participants judged understand speech and being mobile as being

most important (Fig. 6). Surprisingly speaking was rated far less important than understanding speech. Maybe the participants are not aware of the importance of feedback. They might as well not be used to speech as a feedback channel for a robot system because they are used to output modalities like screens.

After the interaction the participants rated BIRON's performance regarding the same abilities. Mobility was excluded because the robot could not move during the experiments. Even though the ratings cannot be compared directly, a dissonance between importance and performance of abilities is found (Fig. 6). Only the voice output of the system excelled the users expectations. Participants stated that BIRON's ability to "understand speech" was poor. They did not know that speech was typed in. When utterances are typed in manually, speech recognition naturally cannot be the problem. In fact, the robot did not understand the users because its vocabulary is limited and it cannot cope with certain sentence structures (e.g. one word sentences do not contain enough information / context for the robot to understand them). Thus, even though this is not the case, users seem to attribute all misunderstandings to speech recognition problems, which again determines changes in their behavior that are described in depth later in this section.

It is noticeable that the standard deviation (SD) for the rating of the importance is higher than for the ratings of the robot BIRON for all items but recognize people. This finding implies that before the interaction users had manifold notions of household robots, their tasks and abilities. In contrast, after the interaction the participants seemed to have formed models of the robot's capabilities. Especially the standard deviation (SD) for speaking and understanding speech is smaller than before (1.2 vs. 0.8 and 1.1 vs. 0.7, respectively). Contrary to this, "recognize people" was rated between 2 and 5 after the interaction. Thus, SD is higher for the rating of BIRON (SD=1.0) than for the importance rating of the item (SD=0.8). SD for the last two items (recognize gesture, recognize mimic) is smaller for the robot ratings but can't be used as a standard of comparison because 6 out of 15 participants indicated that they were not sure. These findings imply that the robot does not emit enough feedback apart from speech. The only other feedback channel was the camera movement. This modality does not seem to convey enough information to enable the subjects to judge the robots abilities to recognize people, mimic, and gestures.

The felt lack of abilities of BIRON probably influenced the users' rating of how much fun they had interacting with the robot (mean 2.7 on a scale of 1 to 5). In general participants cannot imagine to have BIRON at home (mean 1.8 on a scale of 1 to 5). This finding is strongly correlated to the fact that participants cannot imagine to use a domestic robot at all.

Regarding the question how much BIRON's answers helped them to understand what the robot did, a mean of 2.5 on a scale of 1 to 5 was computed. Nevertheless, the standard deviation is very high (SD=1.25), indicating that the feedback helped some participants and did not help others at all. This implies that the feedback of the robot should adapt to each interaction partner. No matter how subjects judged the feedback, all but one indicated that they changed their behavior during the interaction in order to improve it. With a mean of 4.1 (scale 1 to 5) participants stated, that they consciously adapted their behavior to BIRON. The subjects themselves described different adaptation strategies they applied (see Table 1).

All but two subjects only mentioned conscious linguistic adaptations. Participants did not consciously adapt their gestures to the robot, even though a notable number of gesture behaviors are found and changes between them occur as often as linguistic changes. This finding implies that the users only consciously adapt to a modality when it provides feedback. As BIRON has no capabilities to express gestures, the changes in user behavior seem to be an unconscious variation instead of an adaptation. There is no way for users to find out which behavior is beneficial for the robot.

TABLE 1 ADAPTATION BEHAVIORS

Adaptation behavior	**#**
Manner of speaking	8
• speak more clearly	2
• vary intonation	1
• vary loudness	2
• speak more slowly	3
Sentence structure	17
• verification questions	4
• switch between different sentence structures	1
• simple sentences	2
• imperative sentences	2
• one-word-sentences	6
• special sentence structure (“This is a…”)	2
Content of utterances	7
• repetitions	3
• paraphrasing	1
• descriptions	3
Change between gestures	2
• hold object into the camera focus	1
• change between moving object and holding it still	1

= number of participants that mentioned the adaptation behavior

The most common conscious adaptation was to use one-word sentences (sentences that only contain the name of the object shown to the robot). This finding implies that users think that BIRON only understands very simple sentences. Another common behavior is to ask verification questions. These questions show the users' need for more feedback. The results of the video analysis (section B) will shed some more light on the question of changing behavior and adaptation.

Moreover, the subjects had the feeling that they only used a minor percentage of the robot's abilities (mean 40%, max. 60%).

Apart from these general questions about the interaction, participants were asked to rate some items regarding speech. These questions are important since, as described above, speech is the main modality BIRON used in the experiments.

Altogether, ratings for BIRON's voice output were rather high (Fig. 7), especially for intelligibility. With regard to the importance of the modality speech as such this is a very positive result. It indicates that the lack of feedback is rather a question of dialog design than of technological problems.

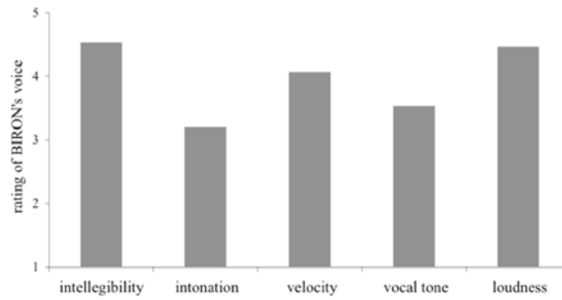

Fig. 7 Rating of voice output

B. Results of the video analysis

Based on the eight linguistic behaviors presented in section III-C, the data implies four situations in which users most frequently change their behavior during the interaction. Most changes occur when BIRON says it has not understood or it cannot do something (1). When this happens, subjects try to paraphrase. Thus, they switch between behavior 1 and 2 (saying a whole sentence, saying one word or a very short phrase). Another important reason for changing behavior is the need to verify if BIRON has understood something (2). This happens when the robot signals that it understood and the user wants to be sure if this is true. Thus, participants ask for feedback (behavior 4), knowledge and abilities (behavior 5) of the system. Another situation that causes users to switch between behaviors is a missing reaction by the robot (3). When BIRON hasn‘t done anything for some time subjects start naming the object in a detailed manner (behavior 1) or describing the object (behavior 3). Last but not least, participants change their behavior when they show a new object to the robot (4). In this case they usually ask BIRON for attention (behavior 8) and name the object in a whole sentence (behavior 1).

The data was also analyzed regarding gestures. With the help of the nine groups described above, five typical situations during which the users switched between different gestures are proposed. Primarily, participants apply another behavior when a new object is chosen (1). Usually the object is presented to the robot (behavior 1), the object is moved in front of the robot (behavior 2, 3), or the subjects point at the object (behavior 7). All these behaviors seem to be applied to gain the robots attention. Thus, in this situation the gestures seem to have the same function as the speech, where behavior 8 (ask BIRON for attention) is most common.

The exact similar behaviors occur when users try to present the same object one more time because BIRON has not recognized it or hasn't done anything in quite some time (2).

As described above, when BIRON has not understood something people paraphrase. While doing this, they as well try two new gestures (3). First of all, subjects hold the objects (behavior 9), which often seems to be a sign of disappointment. Some chose the opposite behavior though and try to regain BIRON's attention by moving the object to another position (behavior 2). This might be due to the fact that people belief, BIRON might not see the object at the current location. The same new behaviors are chosen when BIRON has not done anything for quite some time (4).

The last situation, that typically causes a change in behavior, is the description of an action (5) (e.g.: "This is a pencil. It is used for writing."). In this case a very close coherence of speech and gestures can be seen because the described actions are accompanied by the imitation of the action.

Most common switches of gestures occur between presenting the object and moving it to another position. Thus, there is a constant change between holding the object still for the robot to recognize and trying to obtain the robot's attention.

Altogether, it is noticeable that changes in gestures and speech often occur at the same time and seem to be closely related. This appears natural but is surprising in a way since BIRON only gives feedback with the help of speech and camera movement but does not use gestures. As described above, participants reported almost no conscious changes in gestures. Therefore, changes in gestures can be regarded as variations accompanying speech.

V. Conclusion

In this paper situations were identified in which changes in discursive behavior (gesture and speech) in HRI occur. Knowledge about them is important for robot design, since it is the basis for building models of human behavior which help the robot to recognize what the human is doing or what intention she has. Moreover, it allows for adaptation of the system, which improves the interaction. It became clear, that the feedback of the robot has a major influence on subjects' behavior. Changes in behavior seem to occur consciously when feedback for a certain channel is available. Thus, in our experiments mainly changes in speech were reported by the participants. Moreover, it was found that feedback influenced the attitudes of the users during the interaction. Their views of the speech in- and output of the robot were rather consistent after the experiment, whereas users were not sure about BIRON's abilities to recognize people, mimic and gestures.

Concerning verbal feedback of the robot, it was found that it did not have the same effect on all users. While it was very helpful for some, it did not help others at all. This finding supports the usefulness of adaptive verbal feedback by the robot.

References

[1] S. Oviatt, J. Bernard, and G.-A. Levow, "Linguistic Adaptation During Spoken and Multimodal Erros Resolution", in *Language and Speech*, 1998, 41 [3-4], pp. 419-442.

[2] D. Bohus and A. Rudnicki, "Sorry, I Didn't Catch That! – An Investigation of Nonunderstanding Errors and Recovery Strategies", in *SIGdial-2005*, Lisbon, Portugal.

[3] M. Swerts, D. Litman, and J. Hirschberg, "Corrections in spoken dialogue systems", *in Proc. of the International Conference on Spoken Language*, Bejing, China, 2000.

[4] J. Allwood, "Dialog Coding – Function and Grammar. Göteborg Coding Schemas", Gothenburg Papers in Theoretical Linguistics, 2001.

[5] S. Lee and S. Kiesler, *"Human Mental Models of Humanoid Robots",* International Conference on Robotics and Automation, ICRA 2005, Barcelona, Spain.

[6] K. Fischer and R. Moratz, "From Communicative Strategies to Cognitive Modeling*", in Proc. of the First International Workshop on `Epigenetic Robotics'*, Lund, Sweden, 2001.

[7] S. Kiesler and J. Goetz, "Mental Models and Cooperation with Robotic Assistants*", in Proc. of Computer-Human Interaction (CHI)*, 2002.

[8] A. Powers & S. Kiesler. "The Advisor Robot: Tracing People's Mental Model from a Robot's Physical Attributes", *in Proc. Human-Robot Interaction Conference (HRI 2006),* 2006.

[9] K. Fischer and M. Lohse, Shaping Naive Users' Models of Robots' Situation Awareness, *in Proc. of RO-MAN'07*, Jeju, Korea, 2007.

[10] H. Sacks, Turn-taking. Part 2. Lecture 3. in*: Lectures on Conversation*, Volumes I & II. Malden Oxford Carlton: Blackwell Publishing, 2005 pp. 32-43.

[11] C. L. Breazeal, Regulation and entrainment for human-robot interaction, in D. Rus, & S. Singh (Eds.), *International Journal of Experimental Robotics*, 21(10-11), 2002, pp. 883-902.

[12] C. L. Breazeal, "Toward sociable robots", *Robotics and Autonomous Systems 42*, 2003, pp. 167-175.

[13] A. Haasch, S. Hohenner, S. Hüwel, M. Kleinehagenbrock, S. Lang, I. Toptsis, G. A. Fink, J. Fritsch, B. Wrede, and G. Sagerer, "BIRON – The Bielefeld Robot Companion", *in Proc. Int. Workshop on Advances in Service Robotics*, E. Prassler, G. Lawitzky, P. Fiorini, and M. Hägele (Eds.), pp. 27–32, Stuttgart, Germany, (2004), Fraunhofer IRB.