Involving Users in the Design of a Mobile Office Robot

Helge Hüttenrauch, Anders Green, Mikael Norman, Lars Oestreicher, and Kerstin Severinson Eklundh

***Abstract*—This paper describes the experiences from the iterative design of a fetch-and-carry-robot, to be used by motion-impaired people in an office environment. A user-centered approach was chosen, involving several steps of information elicitation to inform the design. We describe the main elements of the design process, the communication and interaction components of the final prototype system, and an evaluation of the system in the form of a longitudinal study. Results from this study confirmed that continuous testing with users is extremely important in the design process for service robots. The trials have also revealed that interaction design for robots should not focus only on the individual user, but that other members in the environment can be seen "secondary users" or "bystanders" who tend to relate to the robot actively in various ways. We conclude that these social and collaborative issues should be studied in future research.**

***Index Terms*—Human factors, human–robot interaction, man-machine systems, mobile robots.**

I. Introduction

The ADVANCES in robotic technology have implied that service robots are becoming a part of a realistic user scenario, supporting people in their daily activities. There are already commercial robots available for highly specialized tasks, and more advanced integrated systems are being tested in research laboratories. It is a challenge for research in human–robot interaction to bring these robots into the environments of ordinary people, and to elicit information about their needs and activities, to inform the design of usable robot systems.

This paper describes the work of designing a robot to assist users with everyday tasks such as, to fetch and deliver objects in an office environment. The targeted users suffer from physical impairments that make it difficult to move around and to carry objects. One important aim of the project has been to provide an intuitive interface, which a nonexpert user could operate after a short introduction. In addition, we studied the use of the robot, aiming to validate the design, on a long-term basis in a realistic setting.

We have used a combination of methods to guide the design of Cero robot (see Fig. 1). The design process has been iterative, applying results from user trials to inform the design of a prototype or component, which was then used in new tests, and so on, refining the design of the robot at each iteration. We have also conducted a three-month field study to evaluate the robot in an office environment.

Fig. 1. Service robot Cero.

At an early stage, we performed a questionnaire study to assess people's attitudes toward ISRs. Then we used a framework for task analysis, adapted for human–robot interaction, to find out about the users' needs and work tasks in our particular domain. The findings of the survey and the task analysis were used to guide the development of a prototype robot and interface components, which were then used in a simulation study with a small group of users. The combined observations of these user trials were used to inform the overall physical design of the robot as well as the graphical user interface (GUI) and the natural language (NL) interface.

Below, we will describe these methods, and how they have been used in the design of the current robot prototype. Finally we will describe the setup and results of the field study.

Manuscript received September 17, 2002; revised February 2, 2003 and September 22, 2003. This work was supported in part by the Swedish Labour Market Board (AMS), in part by the Swedish Foundation for Strategic Research (SSF), and in part by the Swedish Agency for Innovation Systems (Vinnova). This paper was recommended by Guest Editors R. R. Murphy and E. Rogers.

The authors are with the Interaction and Presentation Laboratory, Department of Numerical Analysis and Computer Science, Royal Institute of Technology (KTH), S-100-44, Stockholm, Sweden (e-mail: hehu@nada.kth.se).

Digital Object Identifier 10.1109/TSMCC.2004.826281

II. Related Research

Service robotics is a wide and fast moving field, where different disciplines and technologies come together. The existing research reflects these diverse approaches, focusing on various aspects of robotic components and/or applications. However, it is striking that little research has been targeted at user-centered design of mobile service robots.

Artificial intelligence research was among the first fields to deal with the issues of machines, intelligence and human users. For example the social implications of intelligent machines were discussed in the 1970s [1], and robots were named as possible applications. However, widespread usage of robots was only vaguely foreseen in the future, and it was predicted that these intelligent machines would reach widespread usage around the year 2000.

Exploring "intelligent systems" for dynamic environments in the late 1980s, Flynn and Brooks [2] pointed toward the importance of developing and testing embodied robots in real environments to identify unanticipated research issues in system design and development. The robots used in these experiments differed from industrial robots, as they were to be perceived as acting intelligently and not only to represent simple machines of automation in production lines. Researchers were not concerned with the actual usage of a function or a service, but rather the aspects of intelligence needed due to the embodiment in physical environments.

The hospital delivery robot HelpMate,[1] closely associated with the work of Joe Engelberger [3] is regarded as the first service robot actually sold commercially and used for delivery services in hospitals. Evans [4] summarizes the HelpMate development and reports about practical experiences made during different implementations. That the HelpMate robotic system could both reduce costs and provide an improved hospital delivery service has been verified experimentally [5]. HelpMate thus, can be seen as a typical service robot, which had its origin in the technology of industrial robots, but was tailored to answer to the requirements of service robots, including the shared workspace of humans and robots, and the operation by nonexpert users.

The beginning diversification of service robots, as separated from industrial robots, was described by Engelhard and Edwards [6], who reviewed the history of service robot technology and its relation to a number of technological domains and research disciplines. Accordingly, service robotics today contains many different service areas and is directed toward a variety of users. This makes it necessary to consider and evaluate various categories of robots and their intended purpose with regard to the technical components used, possible human–machine interfaces, and physical design.

Mobile robots like Rhino, Xavier, MINERVA, or the Care-O-bot offsprings [7]–[10] have guided and entertained museum visitors in public spaces. These robots have been encountered or "used" by a great number of people outside the robot research community. While impressive in themselves, these robots were often limited with respect to their communication and interaction with users (i.e., visitors of the museums).

[1]HelpMate is commercially available from Pyrix Corporation.

Some "public" robots have allowed to be operated by teleoperation, e.g., via a Web interface [8]. Generally, letting a robot only act in a remote location excludes users from experiencing the robot in its physical manifestation. We will assume in the following that a major characteristic of service robots is their ability to also operate in the proximity of users, and to interact closely with them.

A special field of robotics has targeted the assistive or rehabilitation needs of elderly or impaired users. Different types of assistive devices such as, fixed robot workstations, enhanced wheelchairs, wheelchair-mounted manipulators, or other mobile robotic units, have been developed as well as experimentally and clinically validated in a wide variety of systems and projects [11]–[13]. Workers in this field have a tradition of working closely together with potential users and/or their representative organizations. Accordingly we find the most clearly articulated statements for *usability* as part of the requirement specifications in these robotics projects.

Usability as defined in the ISO standard 9241 states how to specify and measure the usability of products. Despite its origin in office work with visual display terminals, the standard's definition of usability is general enough to be helpful even in service robotics. According to the standard it is necessary to identify goals to be achieved with the use of an artifact. Dependent upon the context (including the specified user group, a task to be performed, the equipment available, and the environment of usage), usability is measured with respect to the *effectiveness, efficiency, and satisfaction* with which the users' intended goals can be achieved.

With respect to robotic systems, Teti *et al.* [14] proposed that user evaluations are to include a sequence of steps, the identification of tasks to be performed, the trial-set up, the training of users, and finally, the direct collection of users' feedback. It is explicitly stated that the resulting data to be collected in each of these states should include techniques for the compilation of *functional*, *usability*, and *indirect* (subjective) measurements to be cross-related in a later analysis.

In-depth case studies of the use of robotic assistive technology [15], [16] have pointed to the importance of an assessment that combines the personal value attributed and the individual need for assistance. This aggregated "useworthiness" attribute seems to reappear in the investigation of the social, emotional and environmental factors in the design of eldercare technology as reported by Hirsch *et al.* [17].

The orientation toward end-users calls for the involvement of usability experts, as they have the interdisciplinary tools and methodologies to bring together engineering specialists on one-hand and user groups on the other hand. Trying to strengthen users' involvement in the design stage, user-centered design approaches [18] can be applied to service robot development by iteratively refining designs and prototypes, which are based upon results of empirical evaluations in realistic use contexts.

The complexity of robotic technology has probably been a major factor behind the lack of user studies in realistic contexts. That robots are ready to leave the laboratory for real world deployment today is argued by Austin *et al.* [19], who mention the autonomous recharging of robots as one prerequisite. The title of their paper "Mobile Robotics in the Long Term" points to an important issue, robots will increasingly be run over longer periods of time, surpassing the usual short-term experimental sessions often reported in research papers.

Arguments about the future need for service robots are often made based on demographic reports (see, e.g., [20] for Sweden), pointing toward a sharp increase of the average age in many highly industrialized societies within the next 10–20 years. This is expected to create a growing need for assistance and an explosion in health care costs. Large scale robotic research efforts such as the Humanoid Robot Project (HRP) with one of its applications focusing on "human care" in Japan [21], or the MORPHA scenario of a "robot assistant for housekeeping and home care" [22] explicitly build on this motivation, and aim to find answers to this challenge by developing service robots capable of providing such care and assistance.

Despite these impressive ongoing projects and technical advances, knowledge is still scarce with respect to how users in the targeted groups actually interact with service robots in their everyday environment. In our view this requires developing usable prototypes and embedding them in real settings. There is a need for both qualitative and quantitative studies to assess how the technology is interleaved with human activities, which in turn will yield input to the design process. We believe that this also requires an understanding of the social context surrounding the individual robot user [23].

III. Survey

A problem in designing for human–robot interaction is that ordinary people have little or no experience of robots, which might help to form their expectations and guide their interaction with a service robot. As a pilot study a questionnaire survey was performed with the purpose of assessing people's attitudes toward the use of a mobile, intelligent service robot for household tasks [24]. The study included 134 participants with various background and education, and with equal participation of men and women. The questions concerned a range of topics such as the tasks that users want the robot to perform, the preferred physical design of the robots with respect to size, height, color, and visual appearance, as well as safety and privacy issues.

The results showed that a comparatively large proportion (30–50%) of those who wanted help with a household task were positive toward having a robot to help with it. The majority of the participants preferred to view a service robot as a smart appliance, although other perspectives, such as a personal assistant, were also rated as acceptable. When asked about the degree of independence of a robot, the majority of subjects preferred a robot that does only what it has been instructed to do, and does not act independently.

The survey also included questions about the preferred way of communicating with a service robot, including the choice of modality. The results showed that most participants preferred speech (82%) followed by touch screen (63%), gestures (51%), and command language (45%). Altogether, the survey showed that speech is a preferred form of interacting with a household robot, but that several other forms of interaction are acceptable as a complement.

IV. Task Analysis

In traditional software engineering, task analysis (TA) is sometimes used to find the structure and contents of the tasks [25]. The methods used are based on theories about human behavior in work situations, and many methods use structuring and decomposition of tasks for this purpose. In using TA for service robots we aim to capture the following information:

- user's work procedures and tasks;
- physical design requirements: TA has to describe properties of the payloads for the robot, physical constraints on manipulators and body size, sensors etc. so as to provide a good basis for the physical design;
- function allocation and relation between the user's and the robot's parts of the work. Disabled users often want to do as much as possible themselves, and also the technology constrains the degree of autonomy of the robot;
- user's expectations on the task: The perceived quality of goal fulfillment is a large part in the usability measurement for the robot.

To obtain data for the design of a service robot, we have used interviews with end users about their expectations on robots, as well as actual needs. We have also used focus groups, consisting of people with knowledge of the particular domain of interest. In particular, we discussed the overall design and the functionality to be provided with a group of people working with assistive technology as consultants.

V. Early Simulation Study

During the initial stage of the project, we designed a physical prototype of the robot with which we performed a simulation study according to the Wizard-of-Oz framework [26], [27] with a small group of users. During these trials a "wizard" operator, located in a separate room, controlled the behavior of the robot and provided spoken output through the robot's speech synthesis system. A set of tasks was explained and provided on paper for the users. First the user should instruct the robot to transport a magazine to another person (location), then the user was supposed to accompany the robot to a table where there was coffee. The coffee was to be fetched by the user and carried back by the robot to what had been designated as the "office." Finally the robot should be sent back to its recharging station and told to switch to standby mode. These tasks were considered to be representative for the scenarios envisioned during the specification of the project.

The purpose of the study was to investigate how users acted when presented to a service robot and if they showed any systematic patterns during interaction. A simulation can also help discover potential problems and difficulties at an early stage. To our knowledge very few studies of this kind have been performed with respect to human interaction with service robots, and therefore the nature of the study was explorative. The sessions were recorded on video and logs were kept of the spoken interaction in the system.

Although the study included only six users who had never previously been exposed to operating a robot, the analysis of the data gave some important insights. We found that the interaction with a robot is more complex than traditional screen-based

Actor	Utterance / Action	Observation
R_1	*Okay*	
R_2	*I am in Maria's room*	
U_3	*You are not!*	
U_4	*<Laughs;>* *<Turns to R.;>* *<Steps back.>*	
U_5	*Please, two meters left!*	Gesturing with right arm with a sweeping motion to his left. Looking to the left, returning gaze to robot.
R_6	*Okay.*	User looking at robot.
U_7	*Okay.*	Looks at his task list, bends over to pick up the journal from R's transport-tray, looks at his task list again.
U_8	*Thank you.*	
R_9	*Okay.*	

Fig. 2. Transcript of a Wizard-of-Oz interaction.

Fig. 3. Early robot prototype used in simulation study.[2] .

interaction because both robot and user are mobile and share the same physical space. Many problematic situations had to do with communication concerning spatial movement and specification of location.

The users often lacked sufficient feedback on the robot's state and where it was headed. When user commanded the robot to fetch or deliver an object, the interaction tended to be continuous, in the sense that the user monitored the robot during the whole operation. Although the subjects had been told that the robot neither was able to detect a user, nor could interpret users' gestures, several subjects apparently assumed that this was the case and issued gestures aimed at controlling the robot.

An example of this communication can be seen in the video transcript in Fig. 2. "R" stands for the robot, i.e., the robot's utterances and movements as controlled by the human operator/wizard invisible to the users, and "U" is the user of the robot—the numbering denotes the sequential utterances.

Before the dialogue shown in Fig. 2 actually started the user had specified a mission. In the time duration between the specification of the mission and the situation transcribed, no additional interaction took place; the user just accompanied the robot in the direction of the mission's goal-point. The dialogue then starts when the robot states that it has reached its destination ("I am in Maria's room"). Instead of accepting this statement, the user rejects it as the robot had actually not reached Maria's office, and apparently finds this argumentation with the robot amusing. We can now observe that the user turns to a directive dialogue, specifying a new subgoal concerning short distance movements of the robot ("Please, two m left!"). This need for switching between different modes of commanding the robot was taken up by our speech dialogue system design (see below).

During the post-session interviews following the trials, some users mentioned that they lacked a readily visible front of the robot. At the time of the experiment the robot had a neutral, cylindrical shape, like a hatbox, with a loudspeaker as the only distinguishing feature (see Fig. 3).

As a human wizard operated the robot manually, the time delay of the robot's reacting upon given user commands was apparently too long. Actually, users commented in the interviews that the robot sometimes looked "hesitant," "a bit lost" or "afraid." This seems to show that users expected the robot to respond to commands immediately, give instant feedback upon given commands, and perform tasks without errors.

The simulation study also gave us important insights of a methodological nature. The users had only received a short instruction about how to interact with the robot, which led to great variations in their way of approaching it. We also realized that the training of the Wizard must be given more effort, so that he/she will act more consistently. Importantly, adapting the Wizard-of-Oz method to human–robot interaction requires special consideration of the robot's spatial movements as well as the dialogue routines allowed.

VI. Prototype System

The initial user studies led to the conclusion that a combination of a graphical user interface and a spoken language interface was necessary to support efficient interaction with the robot system. The speech interface is thought to be the primary interaction modality when the robot is close to the user, otherwise the user will control and monitor the robot via the graphical interface.

The physical design of the robot body was developed in several iterative and explorative steps. An industrial designer took part in this stage of the project, and shaped a robot prototype

[2]We *did not* experiment with children although this actually might be interesting—this picture was taken during an experimentation break when the robot attracted the interest of a visitor to the lab.

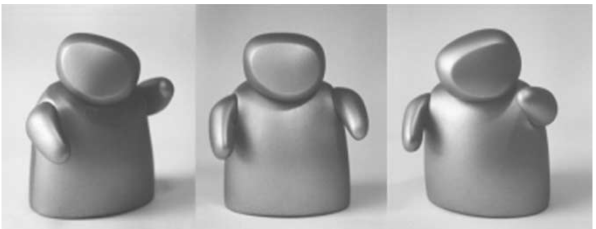

Fig. 4. CERO figure.

based on design discussions and observations from user studies (see Fig. 1). The goal was to fulfill, as far as possible, the different functional requirements on the robot such as the transportation of objects, the system's heat dissipation, maintenance openings, and the placement of interface components. These requirements had to be matched with the constraints set by the robot platform.

As mentioned earlier, we found out in the early simulation trials that the robot needed a more visible front so that users could address the robot more easily. This would help predict the robot's movements and give users the feeling that they were positioned correctly in relation to it. Additionally, we wanted the robot's exterior design and interface to guide how it was perceived, i.e., as natural, intuitive, nonthreatening, and interesting to use.

The need for continuous feedback and a visible front of the robot led us to the development of a life-like character, CERO[3] (see Fig. 4) which is attached to the front of the robot, and which can help to express the robot's state by its movements. This is to be seen as a complement to the spoken dialogue, and a contributing factor for giving the robot a personality. An animated version of the character is also used in the graphical user interface to provide a coherent expression of the visual interface and the robot system as a whole.

As a robot platform, we use a standard Nomadic Super Scout II, which has been enhanced with a transport-compartment as part of the robot body. The robot base is equipped with an industrial PC with a 233 MHz processor and 128 MB of RAM, running Red Hat Linux as the operating system. The onboard computer communicates with the robot's control board driving the motors and taking in the different sensor readings. The robot base has a cylindrical form with a diameter of 0.41 m, and weighs about 23 kg (including batteries). It has a maximum speed of about 1 m/s, a transportation capacity of about 5 kg, and runs about 8 h without recharging. The height of the robot transport compartment is 0.5 m so that objects can be placed on the robot or removed in a sitting posture.

The platform is controlled using a hybrid deliberative/reactive control system [28], which is divided into three layers, a deliberate-, a task-execution, and a reactive layer. The deliberate layer makes intentional decisions according to the state of the system and the requests for missions. The task-execution layer tries to carry out these formulated plans. The reactive layer consists of sensors (e.g., ultrasound) and actuators (e.g., motors) controlled by a number of "behaviors." Behaviors act as tight couplings between the sensors and actuators. For example, an "avoid collision" behavior while the "go-to-point" behavior is active will result in a request to the motor controller to move slower or finally stop, if the ultrasound sensor readings indicate an obstacle in the path of the robot.

[3]Co-operative embodied robot operator, pronounced [sero].

Sixteen ultrasonic sensors are used to find "landmarks" such as edges on furniture, and power outlets in walls etc. Together with the readings from the onboard odometer the landmarks are used to triangulate the robot's position in a predefined map of the environment (see [29] for details).

The human–robot interaction architecture for communication and interaction is organized as a set of distributed components (see Fig. 5) that provide services to each other and interconnect with the I/O devices for the user, e.g., the different graphical interfaces, the CERO figure, or the components of the speech dialog system. Services can either perform tasks for other components or trigger events that are listened to by registered components. The components use common object request broker architecture (CORBA) as the framework for distribution. CORBA provides a high-level distributed object model for systems, as well as being interoperable across different languages, machines, operating systems, and networks and thus, is well suited for the service robot application domain.

VII. Design of the Graphical User Interface

The design of the graphical interface has been kept as simple as possible. To achieve this simplification only a minimum of functions with a high-level feedback about the robot status was aimed for. To guide the design by involving users during the development of the prototype, we conducted a "thinking aloud" walkthrough session of an early version of the GUI. We later applied well-known heuristics [30] and interface design guidelines, e.g., [31], to make the design intuitive and usable. Individual functions, labels, and feedback systems were tested for clarity and improved iteratively.

In the upper left hand corner of the screen (see Fig. 6), a CERO-figure is animated to indicate the robot's status (e.g., "free and ready" or "busy on a mission"). Below is a text field providing feedback about the mission status followed by an emergency switch in the form of a "stop"-button. The content-screen to the right is divided into two major areas. On top is the dynamically updated map of the environment in which the robot can move. The locations to visit are depicted as colored circles in the map. When the user opens the drop-down list and scrolls through the different places available to send the robot to, the location changes to another color (in Fig. 5 a light gray circle in the upper right hand corner indicates the selected "supply cupboard").

Below the map, two basic types of missions can be activated, either a "Go-To"-mission or a "Deliver"-mission. "Go-To" sends the robot from its current position to the place specified, whereas the "Deliver mission" tries to transport an object between two locations. The desired mission option is selected by scrolling in a drop-down menu and then initiated by clicking the "Go"-button. After that, the user can monitor the robot's movement in the graphical map of the environment

Command execution is continuously followed by text-message feedback in the text-field to the left. Possible messages are for example "going to the kitchen," "mission done," or "robot stopped."

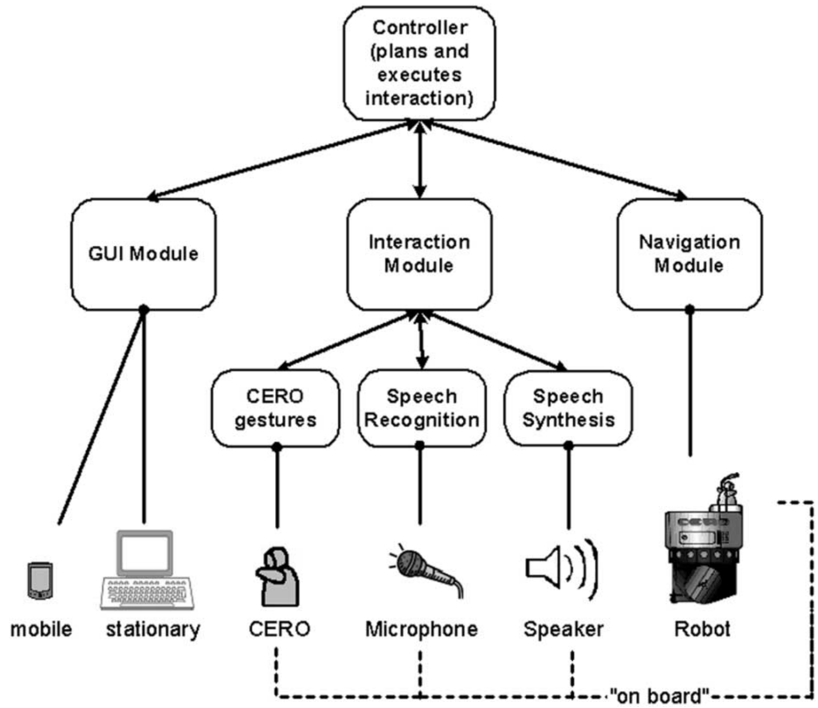

Fig. 5. Communication and interaction components.

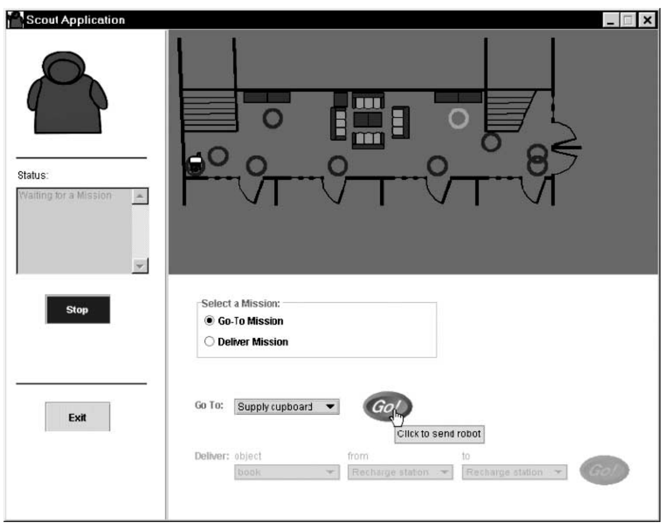

Fig. 6. Screenshot of the graphical user interface.

Commands are structured as imperative sentences, activated with buttons and drop-down selection elements (e.g., "⟨Deliver⟩ ⟨the book⟩ to ⟨Mary's office⟩"). By designing the menus in this fashion we strived to make the phrasing of the speech interface and the command construction in the GUI consistent with each other. In this way, the graphical interface acts as a partial help system for the speech interface, and the coherence of the overall interface is increased.

We found that feedback about the robot system status is extremely important to users, especially when the robot is remotely supervised. The user should therefore, be able to assess the robot's status at any time with only a quick check in the interface. This supervision of the robot is supported by the different representations of the robot status information, including the CERO character, the text field, and the map.

The bird's-eye map of the environment turned out to be an essential, but somewhat problematic part of the interface. Currently, maps need to be manually defined since automatic map generation is still difficult to achieve in our system. Robot expert knowledge is therefore required to produce first the numerical map for the robot navigation system and then its visualization for the user in the graphical user interface. Users of the robot are not likely to be in the position to generate these required system component settings themselves or adopt them to their needs. Furthermore, changes like remodeling of the floor plan at a later point of time are difficult to incorporate as no easy adjustment to the map system is provided for.

In order to investigate this problem, we looked into alternative forms of representation to display to the user the whereabouts of the robot. Fig. 7 shows one of the solutions discussed. The main idea of this sketch was to make the conceptual representation of the robot's location and mission representation independent of the actual physical environment.

The design proposal is using the metaphor of a subway map where lines represent different missions along the robot's routes, and with "stations" signifying goal points. We believe this design has a greater potential of enabling users to setup the robot interface themselves. Scalability of this type of mapping representation is also greater, yielding a more flexible representation. However, so far no user evaluation has been conducted to compare these representations with each other.

VIII. Design of the PDA Interface

We created a portable robot-control interface on (see Fig. 8) a Compaq iPAQ personal digital assistant (PDA) as a complement to the stationary graphical user interface. This mobile de-

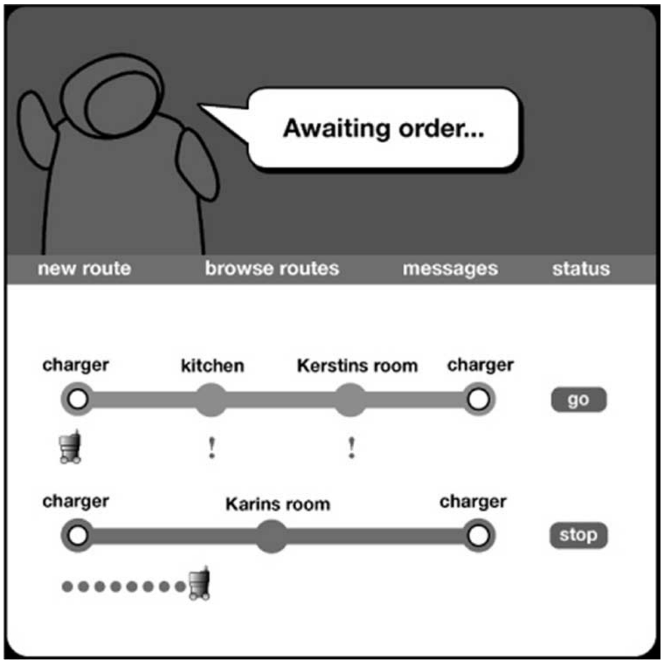

Fig. 7. Robot-missions as subway-lines.

Fig. 8. PDA interface for single hand usage.

vice was meant to provide for mobile and portable robot control while the user is away from her desktop-based robot application. The scenario envisioned that the service robot needed to be instructed while the robot and the user were in close proximity of each other without using the speech interface. The usage of the speech interface should thus become an option in this situation, not necessarily a requirement, giving a user the choice of selecting the modality of communication with the service robot.

For the design of this handheld device we tried different prototype designs and interaction strategies (for details, see [32]). The potential user should for example not be forced to use the device's plastic-stylus as input modality as this would make it necessary to use both hands. Single-hand usage was believed to be a necessary requirement in scenarios where users need crutches as walking aids. However, practical experiences later showed that even the one-handed usage is suboptimal unless a user is sitting down.

An additional limitation was the PDA with its attached network extension card weighed more than 0.4 kg. During the trials this was judged to be too heavy to carry around conveniently. Future mobile service robot controls should instead aim for lighter designs.

IX. Motivation for the Near-Navigation Interface

During the start-up sessions of the longitudinal user study (see below) we identified an issue not anticipated in designing the robot and its interfaces. We had assumed that loading and unloading of the robot would take place while the user is sitting in an (office-) chair. However, the long-term trial was to be conducted in a heritage protected building with high doorsteps, which made traversal into offices impossible for our robot prototype. As a result, the user had to stand up close to the door to place objects on the robot's transport compartment.

This upright posture forced the user to place her crutches in one hand while trying to reach the comparatively low platform with the other (see Fig. 9). As a workaround we developed a "near-navigation" interface on the PDA. Using the joystick button below the screen of the PDA (Fig. 10), the user was able to navigate the robot directly with high precision. This made it possible to drive the robot close to the doorstep and place objects upon the robot sitting in her wheeled office chair.

We believe that important lessons can be learned from this episode of the study. Wheel-based service robots should have the ability to traverse common office environments, including doorsteps, carpets, etc. Given the chance to redesign our robot platform, we would today also propose a height adjustable transport-compartment. Independently of the (un-)loading situations we encountered, it should be possible to reach the transport compartment, either standing, sitting, or perhaps even while lying in a bed.

Generally, we found that users want to control the robot directly in certain situations to adjust its position or direction, and accordingly that the basic mode of specifying mission goal points in the graphical user interface (or with the spoken dialogue interface), was seen as not flexible enough. We thus, learned that both these two modes of navigation (direct control and preprogrammed missions) are essential for the usability of a service robot in this type of use situations. Not only should both modes be provided in the interface, but also (and in particular) the shift from one mode into the other should be carefully designed for to ensure easy and safe operation of the robot.

X. Design of the Spoken Dialogue Interface

Employing a user-centered work model for the speech interface implies a procedure of iterative prototype refinements with increasing complexity. User involvement in this part of the project has so far been limited to participation in simulation studies (see above) and a formative evaluation with a small group of potential users.

A significant part of the design process has concerned consideration of the kind of dialogue to be handled within the system.

Fig. 9. User loading the robot.

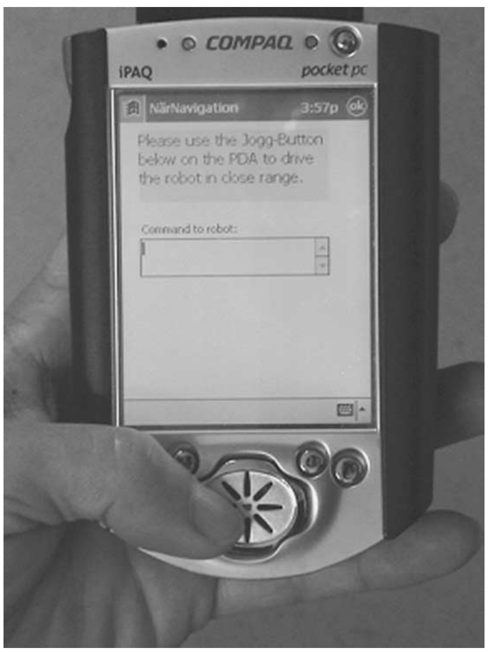

Fig. 10. PDA based near-navigation interface.

Departing from basic tasks that the system can perform, we identified a set of usage scenarios. These scenarios gave a basis for creating a set of relevant example dialogues. Using these *synthetic dialogues* we were able to identify some problematic cases at an early stage without implementing a prototype. Similar methods have been used in other studies as a means of assessing how users envision their interchange with a robot [33], [34]. However, the use of synthetic dialogues can only take us as far as our own educated guesses allow, and therefore, user involvement is necessary to test the proposed design solutions in realistic situations.

Current theories of spoken dialogue can serve as way of achieving naturalness in a human–robot dialogue. According to e.g., Bunt [35] and Allwood [36], participants in a natural dialogue are engaged in a joint co-operative behavior of achieving a common goal. In the dialogue process, grounding, i.e., establishing common knowledge of a dialogue topic is an important way of ensuring and sustaining successful communication (see Clark and Brennan [37] Brennan and Hulteen, [38]). Grounding may concern both the linguistic content and different types of context (Bunt), the latter is becoming increasingly important in robot systems as robots utilize contextual information gathered in their environment to solve tasks. Additional sources of influence for design of dialogue are common practices of speech interface design (see e.g., [39], [40]).

Actor	Utterance	Action
U_1	*Robot, get coffee*	
		[Partially specified command: Get coffee at X]
R_2	*Get coffee, where?*	
		[Request user to specify location X]
U_3	*In the kitchen.*	
		[User specifies X=kitchen]
R_4	*Get coffee in the kitchen?*	
		[Request confirmation]
U_5	*Yes*	*[User confirms]*

Fig. 11. Grounding a command in a synthetic dialogue.

In our system, proper phrasing of the system's contributions is an important part of the process of accumulating common ground in dialogue between the user and the robot. This includes explicit feedback about what the system has perceived as spoken input, as well as questions, prompts etc. that limit the possible range of answers. When a natural language command is received by the system it maps its meaning to goals that are solved by the planning component of the robot. The system uses a consistency check to assure that only fully specified planning constructs are sent to the planner. When the system detects that a command is only partially specified it raises this as an issue by prompting the user for the missing information. When the system has received enough information to perform the task it requests a confirmation of the user before it actually attempts to initiate an execution of the received task. This is illustrated by the dialogue in Fig. 11.

A. Grounding Strategies

In the current dialogue interface for our robot the dialogue can be characterized as primarily task-oriented. In our system we have chosen a careful, or *cautious*, grounding strategy to handle dialogue concerning high-level tasks like "getting an object" or "going to a location." In such dialogues a confirmation is requested for every task involving physical movement of the robot (this is exemplified in Fig. 11). This means that a pending action is treated as grounded and sent to the robot's task planner only if the user explicitly utters a confirmation (e.g., "ok" or "yes, please").

Near-navigation requires another strategy for grounding. The robot is supposed to move only a short distance and if every command was to be explicitly confirmed it would make the process tedious for the user. Dialogues for low-level goals, such

Actor	Utterance / Action
U_1	*Turn left!*
R_2	*<robot turns left>*

Fig. 12. Directive dialogue.

System state	CERO action
On	Not moving
Audio signal received	Raise head
Parsing	Small head nods
Completed	Nod
Parse failed	Shake

Fig. 13. Cero actions and the corresponding system states.

as near navigation, are therefore grounded using an *optimistic* strategy, assuming that commands are grounded immediately when they are received by the system. This is shown in Fig. 12 where the robot immediately starts moving after receiving a turn command.

In order to handle the shift between modes from high-level task specification to low level near-navigation we use a cautious strategy. This means that when a near navigation command is received the user is explicitly prompted to confirm the shift to near-navigation. The system then stays in this near-navigation mode (using an optimistic grounding strategy) as long as it receives commands that are interpreted as near-navigation commands. The shift back to cautious grounding is triggered by two events, either the system receives a high-level task command, e.g., "go to a location," or a time-out occurs when no new commands have been spoken by the user.

XI. Low-Level Feedback Using a Life-Like Character

Concerning low-level grounding, a life-like character intended to produce conversational feedback gestures is used. The development of the CERO (See Fig. 3) concept originated from the Wizard-of-Oz trials where some users commented that the robot seemed very "quiet" when it was neither using its speech synthesis nor moving. Users also informed us that they lacked a sense of direction or heading on the robot.

We interpreted these user responses as a need to provide more explicit *communicative feedback*. In response we devised a life-like character, CERO with the twofold purpose of 1) help providing a visible direction for the robot and 2) providing low-level feedback as a supplement to the spoken feedback issued by the dialogue system.

The movements of the CERO character are designed to be integrated with the speech system so that it is both capable of issuing conversational gestures (e.g., raise or lower its head) reactively, based on system states, and co-expressive conventional gestures (e.g., emblems: nod or shake its head, call for user attention). Some examples of system states and corresponding gestures are shown in Fig. 13.

XII. Long-Term Field Study

Short-term user experiments with an assistive robot only give a limited insight into how a service robot might actually be operated by a user in a real setting, i.e., as the robot becomes familiar and is used frequently in the activities of daily life. With the working robot prototype and its interfaces available, we therefore decided to carry out a study of long-term usage by a user with real assistive needs.

The user participating in the study works in a research department and is partially motion-impaired, i.e., the user walks aided by two crutches. Since both hands are thus occupied, carrying common objects like coffee cups and books is difficult. A secretary to the department staff often needs to help out with transportation tasks, however she is often busy and is off-duty one day a week. The area used for the study is the workplace of the user: an office environment in a public building with two adjoining corridors, providing an area of about 70 m in length.

To keep overall trial complexity low, the robot was to be operated only with the graphical user interface on a desktop computer and a mobile, networked PDA-based interface as a complement. The stationary computer was used from the user's office. As explained earlier, the PDA-based interface provides a portable robot control.

The procedure of utilizing the robot was to have the robot available during regular office hours. The robot would get started up in the morning by a technician; the user herself however was responsible for starting up the Windows PC with the robot application when needed (usually, this was done upon arrival at work in the morning). Once started, the robot waited at its standby location to receive the user's missions.

A. Methodology and Data Collection

The study lasted for three months. This period can be subdivided into three different parts in which different types of data were collected. The trial started off with four introductory sessions over a two-week period where the user was instructed how to use the system for transportation tasks. The user scheduled special time for these sessions, and technical staff and robot researchers were present to help out and to observe interactions directly. Some last-minute adaptations of the system were also done according to observations or issues raised by the user.

After the introductory sessions, when the user had become familiar with the robot and the setup of the trial, she was left with the robot and could use it according to her own needs. Practically speaking, the robot was moved daily to a standby location close to the user's office. The system was started up, and its internal system states monitored for failures, but no special trial sessions were scheduled. The user was not actively encouraged to use the robot other than through its availability. This was thought to allow for a natural and user driven usage pattern to evolve over time.

Data from this second phase was collected by saving the user's voluntary comments in written form to the robot's e-mail address, forming a chronological account of the user's activities during the trial. Additionally, the researchers could observe ongoing robot missions directly. System log files with quantitative data on the robot's missions and the interactions taking place in the interfaces were also collected allowing for later cross-referencing with observations and user's records.

The trial ended with an in-depth interview with the user two weeks after the end of the trial period. The interview covered general questions about the robot usage, clarification of special events observed, reflections on the robot's usefulness as assistive tool, and issues for improving the system.

B. Findings of the Long-Term Study

The long-term study gave a number of important insights and experiences. The findings are mainly qualitative, which is natural since the purpose was to follow the situated long time use of one particular system with one single, committed user. Generalizations from these results should therefore be avoided, yet the observations are important to share as they help validate the design, and may contribute to guide further studies.

During the three-month trial period the robot traveled about 15 km, performing more than 400 transport missions. The largest share of missions conducted were transportations between the kitchen and the user's office, which comprised approximately 40% of all missions performed. The robot's missions normally lasted a few minutes e.g., transporting an object from the user's office to the kitchen was completed in about three minutes. When not in use, the robot waited at its standby location.

During the trial it became evident that using the robot did not keep the user from continuing with the other activities of her work. The service robot only requires the user's full attention during short moments (e.g., 6 s from switching attention to the robot interface to having a mission started). After the initial instruction of the robot, the parallel work activity (meetings, phone calls or writing tasks) was continued and the user's attention was again switched back to the task that had been briefly interrupted.

An interesting observation was made as an effect of this short-term interaction and attention switching. Sometimes the robot was sent on a mission and then seemed to be forgotten by the user. In a few cases this led to a complication: The robot was standing in the middle of the corridor at its destination, making it impossible for cleaning personnel to pass with their cleaning-trolley. The cleaners could not resolve this situation, as they had no possibility of commanding the robot directly. The user was completely unaware of this "robot blocking," working in a remote part of the corridor.

Another frequent observation concerns *secondary robot-users* and *bystanders*. Secondary robot-users refers to people who were not directly involved in the operation of the robot, yet are needed to complete a robot mission successfully. An example of this is the robot mission to fetch coffee at a remote location. As our current robot system has no manipulation system to actually pour a cup with coffee this requires the *collaboration* of other people if the primary robot user does not follow the robot to the coffee machine. The robot must, as a result, contact potential collaborators. Secondary robot users need to be co-operative and aware of the robot's purpose. They must also be permitted to invoke basic functions on the robot, e.g., sending it back once the desired object has been received.

The term *bystander* describes another phenomenon: the robot attracted staff as well as visitors' attention as they spontaneously tried to communicate and interact with it. We often observed people greeting the robot, giving it different nicknames, and making jokes about and with the robot (putting paper hats on it, pretending conversations with it etc.). As the robot generally treats people in its path as obstacles, these interested bystanders made the robot slow down from time to time. It is questionable however, whether the bystanders were even aware of this effect of their action, i.e., preventing the service robot from doing its job.

This attention from other people might be interpreted not only as a general interest in mobile robots or an attractive design of them. It also suggests that the social context is of importance for the behavior of service robots or personal robots. If our observations on this point are valid, the focus on the service robot as a machine or appliance to fulfill a single user's demands needs to be complemented with a more social robot metaphor in functionality, communication, and interaction to incorporate the "secondary" users.

The post-trial interview verified explicitly the benefit of having the robot as transportation aid. The user reported that the robot had been useful and filled a real need, and now that the study was ended it was missed—especially on times when the secretary was off-duty.

Critique was raised in the interview about the nonintegrated interfaces, i.e., the functional split-up between the near-navigation interface on the PDA and the graphical interface on the stationary device. As the near-navigation function suspends the "avoid-obstacle" behavior, this interface differentiation was thought necessary as a safety consideration. Another issue raised was, as already mentioned, that it was felt to be too heavy (about 0.4 kg) to be truly portable and a much simpler, smaller, and lighter "remote control" type of device was wished for.

A general observation often made by interaction researchers and designers is that users will try to push a system to test the limit, and our robot system made no exception. A certain area close to the stairs was defined in the system navigation as a prohibited safety zone for the robot. The user circumvented this imminent safety consideration of the system by using the near-navigation control on the PDA. The reason was that objects needed to be transported from an area below the stairs and therefore the robot was driven as close to the stairs as possible. The observation can be taken as evidence that users over time adapt to and come up with workarounds for possible shortcomings of the system or deliberate safety features. The creativity with which users fulfill their needs with the system is not to be underestimated.

XIII. CONCLUSIONS

Our study has shown that it is both possible and realistic to assist people with the technology we can develop today, in spite of the need for improvement of the system(s) used. Our experiences point toward an increased quality of life as a consequence of the users being more independent with such assistive systems. We would like to encourage more endeavors into the harsh realties of trying real systems with users having real needs. Long-term usage studies of service robots are very important, as they broaden the insights into the research field and give practical guidelines for continued development. Although this study comprised only one user, we have gathered experiences that will be useful for further development and empirical study of robot use.

Among the important lessons learned are that people will want to use a robot as one of many simultaneously ongoing activities. A robot can be left unattended in a corridor for some time or just forgotten. The interface design might take account of this fact and help remind the user, or send the robot back after a certain time.

Furthermore, we observed that there was a need for a "near-navigation" control function to adjust the robot's preprogrammed positions. This might be generalized to provide for both an automated, supervisory type of control, and a direct, manual control means for users. For service robots it seems that both states and a transition between them are typical and need to be addressed.

Finally, our study suggests that service robots are not necessarily a personal device only and human–robot communication and interfaces should be viewed accordingly. The issue of bystanders who want to interact with the robot is only one part of a wider scenario of a workgroup sharing a robot to support transport and information-seeking tasks. Especially in settings with multiple users sharing one robot, or settings where different people encounter the personal service robot, we expect that the need to enable interaction even for "secondary users" will appear, e.g., in work-groups, families or public spaces like airports, hospitals, or exhibitions.

ACKNOWLEDGMENT

The authors would like to acknowledge that E. Espmark created the physical design of the robot and provided help and inspiration during the first stage of the project. They would also like to acknowledge the help from their colleagues at the Centre for Autonomous Systems, CAS, in particular H. Christensen, A. Orebäck and O. Wijk, who allowed for the use of their robot navigation software and provided technical assistance.

REFERENCES

[1] M. A. Boden, "Social implications of intelligent machines," in *Proc. ACM/CSC-ER Annu. Conf.*, pp. 746–775.

[2] A. M. Flynn and R. A. Brooks, "Building robots: expectations and experiences," presented at the *IEEE/RSJ Int. Workshop Intelligent Robots Systems*, Tsukuba, Japan, Sept. 1989.

[3] J. F. Engelberger, *Robotics in Service*. Cambridge, MA: MIT Press, 1989.

[4] J. M. Evans, "HelpMate: an autonomous mobile robot courier for hospitals," in *Proc. IEEE/RSJ/GI Int. Conf. Intell. Robots Systems*, Sept. 12–16, 1994, pp. 1695–1700.

[5] M. D. Rossetti, A. Kumar, and R. A. Felder, "Mobile robot simulation of clinical laboratory deliveries," in *Proc. 30th Conf. Winter Simulation*, 1998, pp. 1415–1422.

[6] K. G. Engelhart and R. A. Edwards, "Human-robot integration for service robotics," in *Human-Robot Interaction*, M. Rahimi and M. Karwowski, Eds. New York: Taylor & Francis, 1992.

[7] W. Burgard, A. Cremers, D. Fox, D. Hahnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun, "The interactive museum tour-guide robot," in *Proc. Nat. Conf. Artificial Intelligence*, Madison, WI, 1998.

[8] R. Simmons, J. Fernandez, R. Goodwin, S. Koenig, and J. O'Sullivan, "Lessons learned from Xavier," *Robot. Autom. Mag.*, vol. 7, no. 1, pp. 33–39, 2000.

[9] J. Thrun, J. Bennewitz, J. Burgard, J. Cremers, J. Dellaert, J. Fox, J. Haehnel, J. Rosenberg, J. Roy, J. Schulte, and J. Schulz, "MINERVA: a second generation mobile tour-guide robot," in *Proc. IEEE Int. Conf. Robotics Automation*, Detroit, MI, May 1999.

[10] B. Graf and M. Hägele, "Dependable interaction with an intelligent home care robot," in *Proc. Proc. IEEE Int. Conf. Robotics Automation* , Seoul, Korea, May 2001, pp. IV–2.

[11] T. Jones, "RAID—toward greater independence in the office & home environment," in *Proc. 6th Int. Conf. Rehabilitation Robotics*, Stanford, CA, 1999, pp. 201–206.

[12] C. Martens, N. Ruchel, O. Lang, O. Ivlev, and A. Gräser, "A FRIEND for assisting handicapped people," *IEEE Robot. Automat. Mag.*, vol. 8, pp. 57–65, Mar. 2001.

[13] S. MacNamara and G. Lacey, "A smart walker for the frail visually impaired," in *Proc. IEEE Int. Conf. Robotics Automation*, San Francisco, CA, Apr. 2000, pp. 1354–1359.

[14] G. Teti, C. Laschi, E. Guglielmelli, C. Mazzà, S. Perini, P. Dario, and M. C. Carrozza, "An experimental method for user-driven design of a modular aid system," in *Proc. 7th Int. Conf. Rehabilitation Robotics*. Paris, France, 2001.

[15] H. Neveryd, H. Eftring, and G. Bolmsjö, "The Swedish experience of rehabilitation robotics," in *Proc. Rehabilitation Robotics Workshop*, Hoensbroek, The Netherlands, 1999, pp. 49–54.

[16] H. Eftring, "The Useworthiness of robots for people with physical disabilities," Ph.D. dissertation, Certec, Lund Univ. , Lund, Sweden, 1999.

[17] T. Hirsch, J. Forlizzi, E. Hyder, J. Goetz, C. Kurtz, and J. Stroback, "The ELDer project: social, emotional, and environmental factors in the design of eldercare technologies," in *Proc. Conf. Universal Usability*. New York: ACM Press, 2000, pp. 72–79.

[18] J. Gulliksen and B. Göransson, "Reengineering the system development process for user centred design," in *Proc. INTERACT 2001*, Tokyo, Japan, July 2001, pp. 359–366.

[19] D. Austin, L. Fletcher, and A. Zelinsky, "Mobile robotics in the long term—exploring the fourth dimension," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots Systems*, Maui, HI, 2001.

[20] Socialdepartmentet, Bilaga 8 till Långtidsutredningen 1999/2000: Kommer det att finnas en hjälpande hand, Fakta Info Direkt, Stockholm, 2000.

[21] H. Inoue, S. Tachi, Y. Nakamura, K. Hirai, N. Ohyu, S. Hirai, K. Tanie, K. Yokoi, and H. Hirukawa, "Overview of humanoid robotics project of METI," in *Proc. 32nd Int. Symp. Robotics*, 2001, pp. 1478–1482.

[22] K. Lay, E. Prassler, R. Dillmann, G. Grunwald, M. Hägele, G. Lawizky, A. Stopp, and W. von Seelen, "MORPHA: communication and interaction with intelligent, anthropomorphic robot assistants," in *Proc. Tagungsband Statustage Leitprojekte Mensch-Technik-Interaktion Wissensgesellschaft*, Saarbrücken, Germany, Oct. 2001.

[23] K.Severinson Eklundh, A. Green, and H. Hüttenrauch, "Social and collaborative aspects of interaction with a service robot," *Robot. Autonom. Syst.*, vol. 42, no. 3–4.

[24] Z. Kahn, "Attitudes Toward Intelligent Service Robots," Royal Institute of Technology, Stockholm, Sweden, Tech. Rep. TRITA-NA-E98421-IPLab-154, 1998.

[25] H. Luczak, "Task analysis," in *Handbook of Human Factors*, G. Salvendy, Ed. New York: Wiley, pp. 341–416.

[26] D. Maulsby, S. Greenberg, and R. Mander, "Prototyping an intelligent agent through Wizard of Oz," in *Proc. Conf. Human Factors Computing Systems*, Amsterdam, The Netherlands, 1993, pp. 277–284.

[27] N. Dahlbäck, A. Jönsson, and L. Ahrenberg, "Wizard of Oz studies—why and how," *Knowl.-Based Syst.*, vol. 6, no. 4, pp. 258–266, 1993.

[28] M. Lindström, A. Orebäck, and H. I. Christensen, "BERRA: a research architecture for service robots," in *Proc. IEEE Int. Conf. Robotics Automation*, vol. 4, 2000, pp. 3278–3283.

[29] O. Wijk and H. I. Christensen, "Localization and navigation of a mobile robot using natural point landmarks extracted from sonar data," *Robot. Autonom. Syst.*, pp. 31–42, 1999.

[30] B. Tognazzini, *TOG on Interface*. Reading, MA: Addison-Wesley, 1996.

[31] B. Sun, "JAVA Look and Feel Design Guidelines," Sun Microsystems, Tech. Rep. Version 1.0.2, Dec. 1999.

[32] H. Hüttenrauch and M. Norman, "PocketCERO—mobile interfaces for service robots," in *Proc. Mobile HCI 2001*, M. D. Dunlop and S. A. Brewster, Eds., Lille, France, Sept. 2001.

[33] I. Isendor, "Mänsklig interaktion med autonom servicerobot," Master's thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, Dept. Numerical Analy. Comput. Sci., 1998.

[34] M. C. Torrance, "Natural Communication With Robots," M.S. thesis, Mass. Institute Technol., Cambridge, Dept. Elect. Eng. Comput. Sci., 1994.

[35] H. C. Bunt, "Dynamic interpretation and dialogue theory," in *The Structure of Multimodal Dialogue*, M. Taylor, D. Bouwhuis, and F. Neel, Eds. Amsterdam, The Netherlands: Benjamins , 2000, vol. 2, pp. 139–166.

[36] J. Allwood, J. Nivre, and E. Ahlsén, "On the semantics and pragmatics of linguistic feedback,", Tech. Rep. 64, 1991.

[37] H. H. Clark and S. E. Brennan, "Grounding in communication," in *Perspectives on Socially Shared Cognition*, L. Resnick, J. Levine, and S. Teasley, Eds. Washington, DC: APA, 1991, pp. 127–149.

[38] S. E. Brennan and E. Hulteen, "Interaction and feedback in a spoken language system: a theoretical framework," *Knowl.-Based Syst.*, vol. 8, no. 2–3, pp. 143–151, 1995.

[39] N. Yankelovich, "How do users know what to say?," *ACM Interact.*, vol. 3, no. 6, 1996.

[40] N. O. Bernsen, H. Dybkjaer, and L. Dybkjaer, *Designing Interactive Speech Systems: From First Ideas to User Testing*. New York: Springer-Verlag, 1998.

Helge Hüttenrauch received the M.Sc. degree in computer science and media from the University of Applied Sciences Furtwangen, Furtwangen, Sweden, in 1995, and is currently pursuing the Ph.D. degree in human–machine interaction from the Royal Institute of Technology (KTH), Stockholm, Sweden.

Before joining Graduate School of Human–Machine Interaction, KTH, he was with Ericsson Business Networks, Stockholm, as a Technical Manager and Usability Specialist for Screen-Based Telephony. His research is focused on human–robot interaction and usability of interfaces.

Anders Green received the M.A. degree in computational linguistics from Göteborg University, Göteborg, Sweden, in 1997, and is currently pursuing the Ph.D. degree in human–machine interaction at the Royal Institute of Technology (KTH), in Stockholm, Sweden.

He participates in the Graduate School of Human–Machine Interaction, KTH, and the Graduate School of Language Technology (GSLT), a collaboration between leading centers in language technology in Sweden, hosted by Göteborg University. His research is focused on human–robot interaction using multisensory natural language user interfaces.

Mikael Norman is pursuing the M.Sc. degree in computer science at the Royal Institute of Technology (KTH), Stockholm, Sweden.

He is working as a Human–Robot Interaction Research Engineer at IPLab, KTH.

Lars Oestreicher received the M.Sc. and Ph.L. degrees in computer science from the University of Uppsala, Uppsala, Sweden, in 1987 and 1991, respectively.

He is a Lecturer and Researcher in human–computer interaction and software engineering at Uppsala University, Uppsala, Sweden, and a Researcher at IPLab, the Royal Institute of Technology, Stockholm, Sweden. His research interests are task analysis, and human robot interaction. He headed the Learning Matters research group at Framkom, focusing on e-Learning issues. He has been chair for the Swedish Interdisciplinary Organization for Human–Computer Interaction (STIMDI), and is the Swedish representative of IFIP Technical Committee No 13 on Human–Computer Interaction.

Kerstin Severinson Eklundh received the Ph.D. degree in communication studies from the University of Linköping, Linköping, Sweden, in 1986.

She is a Professor of human–computer interaction at the Royal Institute of Technology (KTH), Stockholm, Sweden. She is head of the Interaction and Presentation Laboratory (IPLab), an interdisciplinary environment for human–computer interaction research and education. One of her current research projects is "Human interaction with intelligent service robots," which is a collaboration with the Centre for Autonomous Systems. Other research areas at IPLab include computer-supported cooperative work, language technology, and computer-assisted writing.