

Transferring Human-Human Interaction Studies to HRI Scenarios in Public Space

Astrid Weiss¹, Nicole Mirnig¹, Roland Buchner¹, Florian Förster¹, and Manfred Tscheligi¹

¹ ICT&S Center, University of Salzburg, Sigmund-Haffner-Gasse 18,
5020 Salzburg, Austria
{firstname.lastname}@sbg.ac.at

Abstract. This paper presents the contextual analysis of the user requirements for a mobile navigation robot in public space. Three human-human interaction studies were conducted in order to gain a holistic understanding of the public space as interaction context for itinerary requests. All three human-human requirement studies were analyzed with respect to retrieve guidelines for human-robot interaction. This empirical work should contribute by: (1) providing recommendations for a communication structure from a communication studies perspective, (2) providing recommendations for navigation principles for human-robot interaction in public space from a socio-psychological and a HRI perspective, and (3) providing recommendations regarding (confounding) contextual variables from an HCI perspective.

Keywords: Human-Robot Interaction, Human-Human Interaction, Public Space, User Study, User Requirement Analysis

1 Introduction

What happens if a robot is dependent on the support of humans to navigate in an unstructured outdoor environment? In this paper, work performed in the project *IURO: Interactive Urban Robot* [1] is presented. The IURO project considers a human-robot interaction (HRI) scenario, in which a robot is injected into a densely populated public place without any previous topological information and has to navigate to a target point, provided only by means of information it gets from passers-by and its proprioception. To realize such a scenario, the IURO project follows a user-centered design approach focusing on the requirements and constraints that have to be taken into account to achieve successful human-robot interaction between naive users and a robotic system in a public setting. Indications that human-robot interaction scenarios can benefit from an early user-centered design approach [2] and ethnographic approaches “in the wild” [3] already exist. These studies underline the relevance of in-depth understanding of humans’ behaviors and the needs for successful long-term HRI in a domestic setting.

But, what is special about interactions between strangers in public? (1) Interaction in public space is not planned in advance, (2) it takes place in an open space, and (3) it

involves prospective co-participants.. The work presented in the following was performed with the goal to identify characteristics and properties relevant for successful human-robot interaction in public environments.

Three human-human interaction (HHI) studies were conducted in order to gain a holistic understanding of the public space as interaction context for itinerary requests. All three HHI studies were analyzed with respect to retrieve guidelines for human-robot interaction. Thus, the data of these studies was analyzed in comparison with an HRI study and relevant HRI literature with the aim to inform the design of human-robot interaction and subsequent user studies. The human-human ways of interaction in public space are assumed to be the gold standard for the most intuitive social interaction in public space and therefore offer the basis for user-centered design in HRI, however, limitations and differences in case of a robotic interaction partner need to be considered.

2 Motivation and Related Work

Human-robot interaction in public space is mainly short-term interaction between the user and the robotic system. Relevant aspects to improve this interaction for a proactive robotic system that asks for the way are (1) communication, (2) spatial arrangement, and (3) contextual/situational influence factors.

2.1 HRI and Communication

Regarding (natural language) *communication* in HRI previous research was mainly focused on analyzing speech acts and turns [4] as well as miscommunication as such [5]. With the underlying research presented in this paper, we try to tackle the problem of the restricted speech capabilities of a robot from a different angle in retrieving those factors that influence the successfulness of a dialog. Therefore, both human-human and human-robot dialogs were analyzed to develop a set of influencing factors that were rated according to the frequency of their occurrence and correlated with the successfulness of the dialogs. The resulting communication structure including guidelines should provide a helpful means for future research in HRI regarding short-term interaction in public space.

2.2 Spatial Arrangement and HRI

Natural interaction in HRI is typically regarded as the actual interaction of input and output [6]. From a socio-psychological view, however, the interaction between humans starts at a much earlier point, including how people select interaction partners and how humans recognize and approach each other, to finally establish an interaction space [7]. Aiming at natural interaction between humans and robots this pre-beginning phase of an interaction has to be taken into account, especially for encounters between unknown entities in public space. Research about the navigation of robots is still strongly

orientated towards the design and study of delivery tasks or movement through domestic settings. Hereby, people are treated as dynamic obstacles, resulting in robots that avoid collision and keep certain distances to avoid unpleasant feelings for humans [8]. Possible ways to select a person for starting an interaction were investigated, as well as how to subsequently approach the person in public space. The results are presented in several guidelines on navigation for natural human-robot interaction in public space.

2.3 Context and HRI

The research on contextual HRI was so far mainly focused on situatedness [9] in terms of the setting in which user studies are conducted, whereby field trials are assumed to provide more insights on natural reactions on robotic agents than lab-based trials [10]. Within the IURO project, contextual information before the actual interaction between the user and the robot is taken into account, by informing the development process of the robotic system and its interaction model. By means of analyzing videos from human-human studies and a human-robot interaction study we tried to identify relevant context factors (in an HCI understanding of contextual/ubiquitous computing [11]) and to set up a context model for situated HRI in public space. This context model should help to inform the development of the IURO robot and the setup of the user studies, in later stages of the project.

2.4. Bridging HRI in Public Space by understanding HHI

Humans are remarkably good at coordinating their actions with each other to achieve outcomes that are difficult or even impossible to achieve by oneself, such as coordinating trajectories in public space and describing routes to unknown places. Such joint actions require coordination at multiple levels. Individuals must not only agree on a plan of action beforehand [12], but they must also continuously adjust their actions to one another to optimize time and space coordination ([13]; [14]). Understanding the mechanisms underlying human-human joint action has become a major goal in cognitive science during the last years [15]. Previous studies have focused on the role of language as a coordination device (e.g. [16]), on the role of shared representations (e.g. [17]), and on the importance of grounding (e.g. [18]). The results of these studies provide important implications, but also constraints for the development of interactive and pro-active robots that perform tasks together with humans as naturally as possible.

Moreover, evidence can be found in literature that humans tend to socially respond to interactive systems [19] and that humans interpret social behavior patterns into animated objects [20]. The conclusions that can be drawn from all these prior findings in various domains and research approaches for the design of a pro-active robot interacting with humans in an unstructured environment are rather limited.

We consider an ethnomethodologically-oriented approach to identify human-human interaction patterns, which can serve as a bridging component in human-robot interaction. Thus, the three HHI studies, which were executed in the requirement-

gathering phase of the IURO project, had the aim to systematically analyze and link HHI data with the results of previous HRI studies in order to meet the requirements of both parties: the human as responder and the robot as asker. Thereby, guidelines should be developed on how to design the interaction with the IURO robot in terms of: (1) communication, (2) spatial arrangements, and (3) contextual/situational influence.

3 Exploring Itinerary Requests in Public Space

In order to investigate the importance of HHI patterns for HRI and thereby derive design guidelines for the IURO robot, three HHI studies were conducted: Study 1 - "Itinerary Requests", Study 2 - "Pedestrian Selection", Study 3 - "Interaction Reasons". Additionally, footage from a field trial with the Autonomous City Explorer (ACE) robot was analyzed to bridge the gap between differences in HHI and HRI.

The HHI studies were based on a scenario, in which the IURO robot is sent to a pharmacy to buy medicine and deliver it to a patient. It is assumed that the IURO robot was instructed to buy the medicine at the "Old Pharmacy" which is located at "Old Market No. 6", in the old town of Salzburg. Figure 1 shows the location where the studies were conducted. Study 1 "Itinerary Requests" started in the street marked with "A". Study 2 "Pedestrian Selection" and Study 3 "Interaction Reasons" had the additional starting point marked with "B", as this spot is more densely crowded. As Figure 1 shows, there are several alternative routes to get to the pharmacy.

To bridge the gap between HHI and HRI we wanted to compare our results with data gathered in an HRI field trial (conducted in September 2008 in the city center of Munich, Germany) with the Autonomous City Explorer Robot (subsequently called the "ACE" study). ACE is a robot that was developed in a nationally funded pilot study of the IURO project (see Figure 2). The ACE robot interacted via gesture and touch screen input and speech and image output, whereas the interaction was controlled by a finite state machine. The interaction started with the robot greeting a pedestrian and the itinerary request. The robot then asked the pedestrian to point in the direction of the designated goal for establishing a reference point for the itinerary request. Then pedestrians could indicate further directions via buttons on the touch screen. During the interaction ACE builds a topological route graph from the presented route information. At the end the robot thanked the pedestrian and followed the route graph (see [21] for further details). In average the interaction lasted 63 seconds (SD: 34.96). The goal for the IURO robot is that speech and gesture are used as input and output.



Fig. 1. The HHI Study Setting.

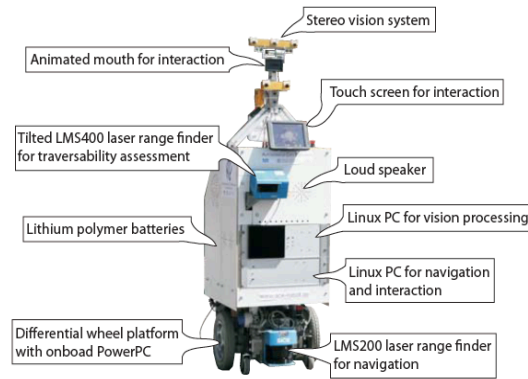


Fig. 2. The “ACE” Study.

The procedure for gathering data on interaction and communication in public space was based on “ecological”, “semi-experimental” encounters, initiated by a researcher or a recruited participant, in all three studies (all studies were conducted in German, which will also be the only language with which the speech processing will work in that specific project). This methodology can be found in various comparable studies (e.g. [21][7]) for a discussion of the difference between “ecologically provoked data” and naturalistic data, see [23].

The data corpus of all three studies is constituted by social encounters in public space in the ancient city centre of Salzburg, Austria (pedestrian area). The encounters were video-recorded (with an informed consent allowing the usage of all collected data) and all studies were supplemented by questionnaires and/or interviews.

The source material (videos and transcripts) of the “Itinerary Requests”, the “Pedestrian Selection”, and the “ACE” study were coded in NVivo¹ by two independent coders using. The predefined coding schemes are explained in more detail in the according section and can be found in [24]. The duration of the footage was about 1 hour for the “ACE” study, 16 minutes for the “Pedestrian Selection” study, and about 4 hours for the “Itinerary Requests” study. Cohen’s Kappa was calculated for the intercoder reliability and all factors with a value above 0.40 were regarded as having sufficient agreement and were thus used for further analyzes [25]. Details on the data analysis and interpretation for the specific areas of interest can be found in the following sections. As the ACE recordings did not offer enough material to explore spatial arrangements, we tried to bridge this gap by HRI literature.

4 Study 1: “Itinerary Requests”

In order to find out more about interpersonal communication in short-term interaction in public space, a participatory observational study was conducted to determine the influencing factors that are crucial for the successfulness of a

¹ www.qsrinternational.com

conversation. These influencing factors were divided into two groups: (1) communication factors – all factors that are inherent in the verbal utterances of the dialog and (2) context factors – all factors that relate to the user or the immediate surroundings of where the dialog takes place. The main focus of the study setting lied in the influential communication factors. To guide our research on the communication factors, the following main research question was addressed:

RQ1.1: What are the important influencing communication factors for successful human-robot conversation in the context of asking for directions in public space?

From a total of 106 pedestrians who were asked during the study, 58 people could correctly describe the way to the “Old Pharmacy”, 3 gave wrong directions (2 of which directed the researcher to a different pharmacy), 21 people did not know the directions to the “Old Pharmacy”, and 18 people did not speak the right language. 26 people did not interact with the researchers at all; the remaining 80 people interacted at a mean interaction time of 19.74 seconds. Those participants, who gave correct directions, interacted with the researchers at an average of 18.02 seconds, whereas the interaction with the participants who gave wrong directions lasted for 22.06 seconds on average (however, the difference is not statistically significant). We could also observe, that 2 people who did not know the pharmacy, tried to explain the way by using a city map or a tourist guide. Furthermore, 2 people were willing to guide the interviewer to the pharmacy. In the end, 41 of the 106 dialogs conformed to our specifications and were included into the following analyses. The results of this study were incorporated in the development of the communication guidelines (see section 7.1) and context influence (see section 7.3). More details on the results can be found in the according sections.

5 Study 2: “Pedestrian Selection”

For humans it is a quite common thing to ask an unknown person for information in public space. However, if there are plenty of people around (as in a crowded city centre), how does one decide which person to approach? Depending on which criteria does the person seeking information select a passer-by. The study aimed at revealing the decisive factors that account for which pedestrian is selected to ask for the way. Although not all preliminary factors derived from this study can be directly transferred into an algorithm for the robot architecture, they will inform the interaction model of how the IURO robot could choose and approach pedestrians. As a second focus of interest the preliminary study was set up to investigate how the chosen people are approached in a crowded area. To support the design of navigation and interaction patterns of the IURO robot, the study investigated the following two research questions for human-human interaction:

RQ2.1: According to which criteria do people select passers-by to ask for directions?

RQ2.2: How is a selected person approached?

The study was conducted from June 10th to 18th, 2010. In total, 20 recruited participants (10 male and 10 female) randomly asked pedestrians for the way to the “Old Pharmacy”. The 20 recruited participants needed to approach 47 pedestrians in total to successfully reach the “Old Pharmacy”, with a minimum of 1 and a maximum of 5 passers-by approached. A total of 2 out of 20 participants did not succeed in finding the requested spot. In both cases the task was stopped after 10 minutes.

To determine the criteria and characteristics people use to identify and select passers-by, interviews were conducted immediately after the destination had been found. All interviews were transcribed and annotated with an ethnomethodological understanding, subsequently the annotations were grouped into categories. However, this quantification of results should not be considered as a one-to-one functional representation, but as a trend or tendency.

In a first step, reasons for approaching a person to ask for directions were extracted. The answers and explanations of the study participants were divided into criteria for selection and exclusion as well as categorized as procedure to select a person (an overview on all 11 selection and exclusion criteria can be found in [24]). Interestingly, the participants more often named persons they actually excluded as a potential source of information (n=20) - with tourist being the most frequently cited group - than criteria for selected persons (n=10). The reason for this is probably that the old town of Salzburg is crowded with tourists, accounting for most of the passers-by.

As the most frequent reasons to choose a person to ask for the itinerary, the participants named “ask a salesman of a shop”. They stated that a salesperson is usually familiar with the surroundings of the shop and very often lives in the town. As second most frequent reason/procedure people named “find some one who is not a tourist” which corresponds with the exclusion reasons. Related to this is “find a local” as an approach. Other approaches were to randomly choose a passer-by, talk to the next person available and to approach a single person.

In a second step, all characteristics of the selected passers-by as named by the study participants were analysed and categorized. In total, 16 characteristics were identified plus a category for exclusion criteria (see [24] for details). The most frequently mentioned characteristic described by the study participants was movement (n=10), followed by behaviour (n=8) and dress style (m=7). 9 characteristics that lead to the exclusion of people were identified. Again the most frequently mentioned characteristic was movement (n=7). Three characteristics referred to objects that people carried with them and which lead to an exclusion: camera, suitcases and a hat.

All interaction recordings were then analyzed with respect to the direction from which the participant approached the passer-by. Four directions were classified according to the empirical findings. From behind; lateral; diagonal frontal and frontal (see Figure 3).

17 of the approached pedestrians were moving towards the study participant and 16 were standing (most of them

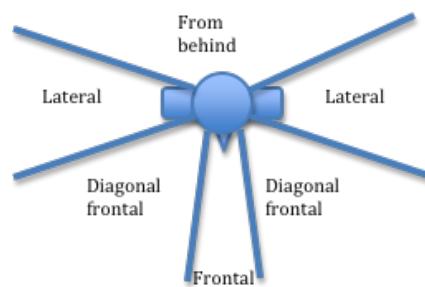


Fig. 3. Categorized approach directions.

were salesmen or waiters in front of their stores or restaurants). Diagonally moving pedestrians and sitting participants were only approached three times, respectively and a person walking the same direction as the study participant was only approached once.

If one considers the movement of the approached person and the direction of approach, several patterns emerge. Passers-by moving towards a participant were most of the time approached from diagonal frontal, only one participant approached in a frontal manner. Standing passers-by, however, were approached from all four directions, no clear preference could be observed. Finally, we observed how the study participants re-directed their trajectories in order to start an interaction with a passer-by. With all the recorded interactions we could identify 3 different types:

- Crossing trajectory: The participants approached the person in a way that both projected trajectories crossed.
- Directly towards the person: The projected trajectory of the participant directly aimed at the selected passer-by.
- Trajectories not crossing: The participant approached in a manner that his/her projected trajectory did not interfere with the one of a passer-by.

Directly walking towards the person could mainly be observed for passer-by who were either standing or sitting or who held a role such as waiter or salesperson. However, some participants also approached passers-by that were moving towards them in a direct way. On the other hand, crossing trajectories as well as trajectories not crossing could almost only be observed when people were walking towards the participant, and the approach took place from diagonal frontal.

6 Study 3: “Interaction Reasons”

Stopping another person in a public space to ask for something is the so-called “opening sequence”, which takes place before the actual start of a social interaction [7]. The mutual spatial arrangement happens in this phase and it concludes with the transition from moving to standing. In this moment the so-called interpellation happens: An unfocused pedestrian turns to a focused “would-be-imminent-co-participant” [26]. To support and maintain the mutual orientation, people tend to give reasons for the encounter with an unknown person as quick as possible [27]. However, it is still an open question to which degree the moment of mentioning the interaction reason influences the itinerary request situation. Thus, the following two main research questions guided this study.

RQ3.1: Does the timing of the reason mentioning influence the interaction?

RQ3.2: Can additional characteristics for pedestrian selection be identified?

The study was conducted in situ, in the old town of Salzburg from June 17th to 19th, 2010. It was based on social encounters, initiated by the researcher. A recruited student (female, 22 years old) took the initiative and asked for directions. She was confronted with the task of selecting the future addressee and organizing the entry of the co-

participant in the interaction in accordance to two different conditions: condition A – subsequent reason, condition B – immediate reason. In condition A, the reason for the encounter was communicated in the pre-beginning before a common interaction space had been established. In condition B, the reason for the encounter was only communicated once an interaction space had been established (i.e. co-participant stopped walking and mutual gaze is established).

In total, the recruited student had to initiate 42 conversations to achieve 10 valid questionnaires. Half of the participants were male and the other half female, their age ranged from 16 to 79 years. The first verbal reaction of the co-participant towards the itinerary request of the recruited student was a confirming feedback statement “yes” (n=4) or a feedback statement combined with a politeness key word “yes, please” (n=4), the other two responses were “hello”. None of these responses were introduced by an “euh” to gain thinking time.

To achieve the corpus, 30 trials were necessary in condition A and only 11 in condition B. However, the retained itinerary descriptions were correct in both conditions only twice. In condition A, two co-participants did not know where the “Old Pharmacy” is located and one co-participant gave a wrong direction. In condition B, a wrong direction was given three times.

The study showed that the timing of the reason mentioning influences the interaction. Condition B was the more appealing one for the co-participants. On a retrospective level it was clear for all co-participants in condition B that the reason for stopping them was an itinerary request. In condition A, two co-participants estimated “needing help” and “searching for something” as reason. Only in condition B one co-participant noticed that the recruited student is going to need something before the conversation started and mentioned the “inquiring gaze” as indicator.

All co-participants of condition B retrospectively had the feeling that they were a conversation partner directly after the introduction words “Euh, excuse me”. This indicates that the follow up sentence “I need your help” (in condition A) did not foster the transition to a co-participant more than the immediate mentioning of the reason while walking. In condition B, all co-participants mentioned “stop walking” as the starting point for the transition. In condition A, this was only mentioned by two participants. In condition B, all co-participants chose almost all response possibilities how they tried to demonstrate their willingness to support the recruited student in her itinerary request: (1) stopping, (2) approaching, (3) answering, (4) smiling, (5) eye contact, (6) turning towards the asker. In condition A, only answer 1 and 3 was chosen, which indicates that the immediate reason mentioning in condition B supports the willingness of passers-by to become co-participants. Interestingly, the mentioned reason why co-participants wanted to help was the same in both conditions: the co-participants wanted to be polite.

But upon what reasons did the recruited student select the pedestrians? Regarding the degree of difficulty for the recruited participant, the reflection form revealed that it was easy to select someone, if there were many people, that it was mediocre if there were only few people, and that it was difficult if there were many tourists frequenting the public space.

Regarding the reasons why a specific pedestrian was chosen as an interaction partner by the recruited student, it could be revealed that she frequently changed her selection strategies during the study and tried to ask younger people, older people,

male, female etc. Three main cues on which the recruited student mainly focused for the pedestrian selection could be identified: (1) pedestrians who looked friendly, (2) pedestrians who walked goal-oriented, (3) item – signalling that the person is a local.

7 Findings for HRI in Public Space

In order to develop guidelines for HRI in public space, we linked the HHI data described in the previous sections with data of the “ACE” study and related HRI literature with the aim to consider both parties during the itinerary request, namely the human and the robot. This data analysis leads to (1) a communication structure, (2) principles on spatial arrangements, and (3) a context model for HRI in public space with the IURO robot, as described in the following.

7.1 A Communication Structure for HRI in Public Space

For the development of a communication structure for itinerary requests in public space, we first analyzed the dialogues of study 1 “Itinerary Requests” and subsequently the data of the “ACE” study. The results from both data sets taken together should help to get a more profound communication structure that pays respect to both interaction partners – the human and the robot.

The following procedure was chosen to analyze the data: First, a coding scheme including the relevant factors that possibly influence the conversational process was developed. Second, two coders coded both data sets independently according to the coding scheme. Third, the results of the coding were analyzed and summarized in terms of a communication structure including guidelines for setting up successful human-robot itinerary requests.

All circumstances that possibly influence the course of the conversation and that become manifest mainly in the verbal utterances are considered influencing factors for successful communication. The coding scheme was set up as a hierarchy and the coded data was further analyzed with respect to the factors: gender of the researcher, gender of the pedestrian, politeness², as well as some basic statistical attributes.

The guiding communication model underlying the communication structure presented in this section, is the model of Shannon and Weaver [28], as the model describes the communication process from a technical point of view (not considering semantics), including the concept of the noise source as a reason for unsuccessful communication. Starting from the noise source, a communication structure for successful human-robot interaction including guidelines for future research design was established (also not considering semantics as such, but communication as a process of sending and receiving information). All retrieved factors from the HHI dialogs and the HRI footage were correlated with the attribute “successfulness”, in order to identify which factors help to ask for directions effectively and which factors hinder the

² in order to distinguish polite from impolite utterances, we identified three different attributes: (1) salutation (entschuldigung- excuse me), (2) request (bitte-please), and (3) subjunctive (könnten Sie – could you).

positive outcome of the conversation.

- “Feedback” was the most frequently occurring positive influencing factor for the successfulness of a dialog, coded in the human-human dialogs .
- The factors “coherence” and “incoherence” tie in with “feedback” as, according to the coding scheme, they indicate instances where an utterance either matches or does not match the immediately afterwards following action – which is a form of feedback. So, if an instance is coded as “incoherent” it can be said that there was either no feedback, not sufficient feedback, no matching feedback, or feedback that was not timed well enough.
- “Fun” and “insecurity” might help to get a pedestrian to stop and help; “fun” during an interaction could help to enhance the tolerance of the pedestrian towards the robot (e.g. regarding non- or merely partial conformity to human conventions concerning conversation), whereas the absence of fun, curiosity, and tension might lead to an early abortion of the conversation.
- Providing the pedestrian with a map might also help to retrieve the desired directions (however, clearly this approach is not part of the IURO project as it is contradicting with the basic assumption that the IURO robot has no map knowledge).
- People tend to unconsciously treat the autonomously navigating ACE robot like another human being and thus expect their conversational partner to act according to human-human communication conventions.
- Even if the condition “politeness” did not result in significant differences regarding the dialogs’ success rate, it will not hurt to make the robot ask politely.

Concerning the directions coded at the human-human study, the following conclusions can be drawn on what a robot should be able to do from a user-centered perspective. A robot should:

- be able to process verbal directions in terms of what direction to go to, which are potentially completed with a gesture indicating the direction,
- be able to recognize landmarks (identify a traffic light, a church, a fountain, etc.),
- be able to process explicit distances (100 meters) and implicit distances (for a short way),
- be able to interpret confidence in a route description by means of probability weightings of route descriptions.

Summarizing the conclusions on the factors and the robot’s abilities concerning directions, the following communication structure for successful short-term human-robot conversation in public space can be established (see Figure 4):

In itinerary requests in public space the human will use verbal directions, gestures, reference points, context information and distance declarations as input modality.

The robot should provide timed feedback, in a coherent manner, and offer a city map as output modality, in case verbal communication is not successful.

Moreover, if the robot expresses its neediness and the interaction is enjoyable for the pedestrian, it is more likely to be successful.

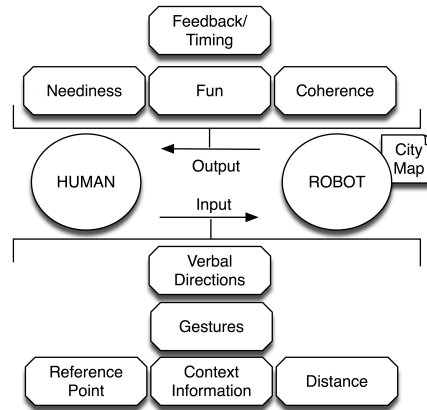


Fig. 4. Communication structure.

7.2 Principles on Spatial Arrangements for HRI in Public Space

Most state-of-the-art HRI studies on navigation are conducted in a context in which the user gets familiar with the robot, e.g. a domestic setting [29]. In such a context the user can adapt to the moving patterns of the robotic platform. In public space however, users will very often be novice, not having made any contact with the IURO robot before. This has to be taken into account when designing the robot's navigation. Usually, the navigation of robots is designed and studied for missions of delivery or simple movement, in which people are treated as "dynamic obstacles" [6]. However, if navigation has the purpose of human-robot interaction, different constraints and requirements have to be taken into consideration, namely the initiation of the interaction, the spatial distance, and the coordination of trajectories.

In general, we can distinguish between three types of "robot navigation": (1) autonomous navigation (moving from A to B), (2) navigation with the aim to start an interaction with a human (approaching a human), (3) navigation during the interaction (adaptation of the interactional space).

The principles on spatial arrangements derived from the studies 2 "Pedestrian Selection" and 3 "Interaction Reasons", mainly address the second and partly the third type of robot navigation (a list of all principles can be found in [24]). In the following we present the concrete recommendations for the IURO robot derived by taking together the results from both studies in combination with findings from the related HRI literature.

- It will be relevant that the IURO robot is equipped with "eyes" (these "eyes" do not need to be the actual cameras of the robot, but need to be cues that are interpreted as eyes by the user) and let these eyes reflect and follow the gaze of the pedestrian.
- It will be relevant that the IURO robot approaches pedestrians lateral or diagonal to initiate an interaction. This approaching direction will also support pedestrians' in

their perception if the robot is (1) heading from A to B, (2) listening to a pedestrian, and (3) noticeable seeking for help (“needy looking”, e.g. when it gets stuck).

- It will be relevant that the IURO robot keeps the ideal spatial distance, whereas it needs to be distinguished between (1) approaching distance (between 0.45 and 1.2 meter) and passing distance (assumed to be less than 1.2 meters). Further user studies will be necessary to investigate this aspect.
- It will be relevant that the IURO robot navigates in different speeds: lower for passing people and faster for approaching/ selecting people and goal navigation (heading from A to B). This will also foster pedestrians’ perception if the IURO robot needs help or is just navigating. Further user studies will be necessary to investigate this aspect.
- It will be helpful that the IURO robot maintains in its position-alignment towards the pedestrian, therefore the user has the feeling that the IURO robot has the same reference point. In other words it will be helpful that the IURO robot keeps the reference point towards the user, therefore they orientate in the same direction during the interaction. Hereby the robot does not need to move the head or the torso, but just move in the same direction as the user, if required.
- It will be helpful, that the IURO robot selects pedestrians depending on their walking speed. As the HH-studies showed, pedestrians who walk “faster than average” and “average speed” are locals, who know the environment and can give valid route descriptions. Slowly, walking people are searching for the route themselves or are non-locals, who are sauntering through the streets. Thus, the IURO robot should select people walking “average speed and faster”. Further user studies will be necessary to verify these speeds.
- It will be helpful to foster the moment of transition from an unknown pedestrian to an IURO interaction partner in the dialogue, by immediately stating the reason for stopping in the itinerary request, namely “I am searching for xy”.

7.3 A Context Model for Situated HRI in Public Space

In order to get information on influencing context factors, the video recordings of three studies were annotated by two independent coders: the “Itinerary Requests” study, the “Pedestrian Selection” study, and the “ACE” study. Before setting up a coding scheme, a model of the human-robot context was developed. Hereby, three entities were considered to be essential for a basic context model:

Robot context factors: Factors which directly affect the robot when interacting with a user or trying to find an interaction partner, e.g. if the user is not able to recognize the content on the screen of the robot because of bright sun light. This means that the factor sun negatively influences the interaction between the user and the robot.

Human context factors: Factors which affect the human during the interaction, e.g. other people watching the interaction, the user being a foreigner, etc.

Environment context factors: Factors which affect the whole interaction between the human and the robot, e.g. cobble stone, narrow space.

For the identification of the relevant contextual influence factors for HRI in public space two coding schemes were developed, one for the videos of the “ACE” study and one for the two human-human studies (“Pedestrian Selection” and “Itinerary Requests”), which both consisted of the same four main categories: action sequence, interaction partner, passer-by context factor, and environment context factor. The passer-by context factor tier should depict which factors appeared during the interaction with the robot. Factors influencing the interaction on the robot’s side could also be found in the environment context factors. Figure 5 depicts the context models derived for the “Itinerary Requests” study and the “Pedestrian Selection” study. The models were developed based on the video annotations of the two independent coders³. The numbers next to the arrows visualize the frequency of annotations. The second model demonstrates the interdependency between user context factors and environment context factors (the intercoder reliabilities were unfortunately too low to see the same effect in the first model). The results from the two human-human studies taken together, the following conclusions can be drawn concerning context factors which are relevant for human-robot interaction in public space.

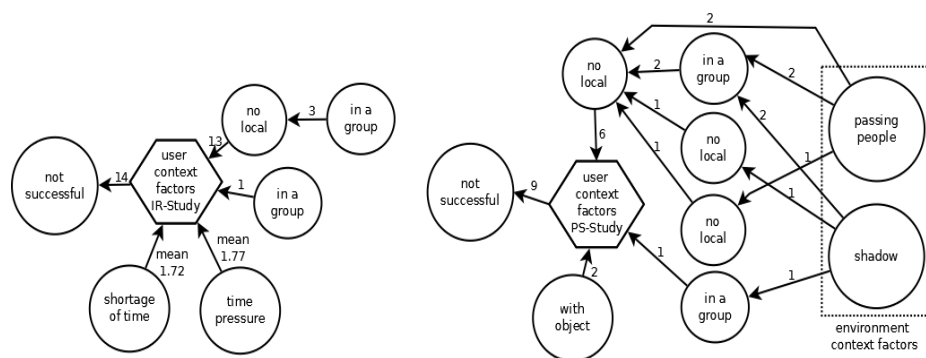


Fig. 5: Context factor models for HHI studies “

User context factors:

- Local: The successfulness of the interaction is highly dependent on the fact if the asked person is a local or not. However, also non-locals were often willing to give advice (e.g. by using their city maps), but their advice was often incorrect, as the location of the “Old Pharmacy” is local knowledge and not included in official tourist guides. Thus, it is important to identify possible locals as interaction partners and to weight the information from a local higher than that of a tourist.
- In a group: There is not much difference in the success rate of the itinerary description provided by pedestrians in a group compared to pedestrians who are on their own. The IURO robot does not need to distinguish between pedestrians walking alone or in groups when it is asking for the way.

³ Only codings with a Cohen’s Kappa value greater than or equal to 0.4 were taken into account.

- With an object: The fact if a pedestrian keeps an object with him/herself during the interaction, e.g. a bicycle, a dog or a shopping bag, does not influence the success rate of the interaction. However, it has to be kept in mind that pedestrians can be limited in their interaction possibilities due to the object.
- Shortage of time and time-pressure: These aspects were only explored in the “Itinerary Requests” study by means of a questionnaire, but both factors revealed to be influential. Pedestrians, who subjectively perceived time-pressure were less willing to help, than people who had an actual shortage of time. However, if the time was too short, the pedestrians did not even stop for a request.

Environment context factors:

- Shadow: The fact if the itinerary request happened in bright sunlight or in a shadowed area influenced the interaction. It could be noticed several times that pedestrians jointly moved to a shadowed area to give an advice. Moreover, in the ACE study people often had to cover the robots’ display with their hands to improve the bad readability, because of the sunlight.
- Passing people: In situations in which many other people walked past, the interaction between the questioner and the respondent was more often successful. This could be due to the fact that “public pressure” forces people to be more supportive during an itinerary request. Itinerary requests more often led to a successful communication if a pedestrian was selected in an environment where other passers-by were present.
- Store: People tried to get information from staff members in local stores in order to enhance the chance to get the correct directions. They were standing in front of the entrance of the store (e.g. a waiter in front of the restaurant entrance) or were owners of outdoor stores (e.g. ice cream vendors). Thus, it is recommended that the IURO robot, which is able to pass low doorsills, prefers asking staff members or storeowners for the way instead of passers-by.

Robot context factors:

Figure 7 visualizes a context model for robot context factors that were derived from the “ACE” study. Again, the numbers next to the arrows show the coding frequencies and only annotations with an intercoder reliability of Cohen’s Kappa greater than or equal to 0.4 were taken into account for the model.

- Moving/not moving: More pedestrians were willing to interact with the IURO robot when it was moving, e.g. approaching people, than when it was standing still. Moreover, the success rate was higher when pedestrians interacted with the robot, which proactively approached them. The longest periods of no interaction were when the robot was stuck somewhere and did not move at all. It is advisable that the robot is somehow moving in such a situation to express its “neediness”.
- Observing/not observing: Pedestrians paid more attention to the robot while it was moving. This implies that it is helpful to make the robot move slightly during the interaction with the user to increase the humans’ attention.

- User: Men were more often willing to interact with the robot in public space. Moreover, their interactions took longer (interestingly in the human-human studies there was no difference in the length of the itinerary dialogues between men and women). Children were also willing to interact with the robot.

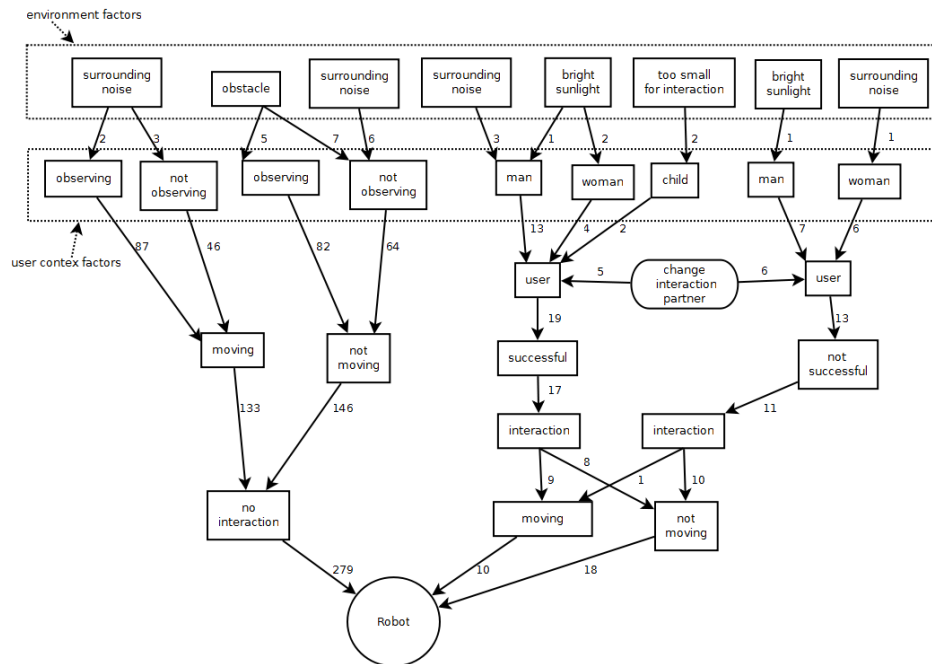


Fig. 7: Context factor model “ACE” Study.

8 Conclusion

In this paper, we presented the systematic analysis of data derived from three human-human observational studies and one human-robot interaction field trial in order to inform the interaction design of a proactive navigation robot in terms of its (1) communication structure, (2) its spatial arrangement principles, and (3) the contextual influence factors. These pilot studies, moreover, have the aim to inform later HRI user studies in the project *IURO: Interactive Urban Robot*. The IURO project considers a human-robot interaction scenario, in which a robot is injected into a densely crowded public place without any previous topological information and has to find its way to pre-defined places, people or items in quickly changing environments through proactive communication with passers-by. All three HHI studies were set up in accordance to the interaction scenario the robot should perform later on in the project, namely an itinerary request in the city center in order to find the “Old Pharmacy”.

We could derive several relevant guidelines and recommendations for the human-robot interaction scenario, whereas some of them need validation through user studies.

Based on the findings presented in this paper and in conjunction with findings on morphology aspects of the robot (considering anthropomorphism and the uncanny valley effect [31] the dialogue model, the global path planning, and the local collision avoidance system for the IURO robot are currently being developed. We are now planning to perform laboratory-based user studies to test the communication structure and the navigation principles on navigation speed and spatial distances in Wizard-of-Oz studies. Successive field trials will be conducted to prove the effectiveness of our strategy to inform the human-robot interaction scenario by means of human-human studies. Moreover, the aim of the successive laboratory-based studies and field trials is to identify key aspects that cannot or should not be directly transferred from human-human interaction to human-robot interaction, as they do not bridge the users' cognitive model about the robot and create wrong expectations about its capabilities.

References

1. <http://www.iuro-project.eu/>
2. Lee, M., Forlizzi, J., Rybski, P. E., Crabbe, F., Chung, W., Finkle, J., Glaser, E., and Kiesler, S.: The snackbot: Documenting the design of a robot for long-term human-robot interaction. In: Proceedings of the 4th ACM/IEEE international Conference on Human Robot interaction, pp. 7--14, New York, NY. ACM. (2009)
3. Sung, J., Christensen, H. and Grinter, R.: Robots in the wild: Understanding long-term use. In: Proceedings of the 4th ACM/IEEE international Conference on Human Robot interaction, pp. 45--52, New York, NY. ACM. (2009)
4. Lauria, S., Bugmann, G., Kyriacou, T. and Klein, E.: Instruction Based Learning: How to instruct a personal robot to find HAL. In: Proceedings of the 9th European Workshop on Learning Robots. (2001)
5. Koulouri, T. and Lauria, S.: A WOz Framework for Exploring Miscommunication in HRI. In: Proceedings of the AISB Symposium on New Frontiers in Human-Robot Interaction. (2009)
6. Sisbot, E. A., Alami, R., Simeon, T., Dautenhahn, K., Walters, W., Woods, S., Koay, K. L. and Nehaniv C.: Navigation in the presence of humans. In: 5th IEEE-RAS International Conference on Humanoid Robots. (2005)
7. Mondada, L.: Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics*, 41, pp. 1977--1997 (2009)
8. Koay, K., Dautenhahn, K., Woods, S. and Walters, M.: Empirical results from using a comfort level device in human-robot interaction studies. In: Proceedings of HRI '06: Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction', pp. 194--201, New York, NY, USA, ACM. (2006)
9. Suchman, L.: *Human-Machine Reconfigurations: Plans and situated actions* (2nd edition). Cambridge University Press. (2007)
10. Sabanovic, S., Michalowski, M. P. and Simmons, R.: Robots in the wild: Observing human-robot social interaction outside the lab. In Proceedings of the 9th International Workshop on Advanced Motion Control (AMC 2006), pp. 596--601. (2006)
11. Dourish, P. What we talk about when we talk about context, In: *Personal and Ubiquitous Computing*, Vol. 8, No. 1, pp. 19--30. (2004)
12. Clark, H. H.: *Language Use*. Cambridge University Press. (1996)
13. Bosga, J., & Meulenbroek, R. G. J.: Joint-action coordination of redundant force contributions in a virtual lifting task. *Motor Control*, 11, pp. 234--257. (2007)

14. Knoblich, G. & Jordan, J. S.: Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(5), pp. 1006--1016. (2003)
15. Galantucci, B., & Sebanz, N.: Joint action: Current perspectives. *Topics in Cognitive Science*, 1, pp. 255--259. (2009)
16. Clark, H.H., & Krych, M.A.: Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50, pp. 62--81. (2004)
17. Sebanz, N., Knoblich, G., & Prinz, W.: How two share a task: Corepresenting stimulus--response mappings. *Journal of Experimental Psychology: Human Perception and Performance*, 31(6), pp. 1234--1246. (2005)
18. Ros, R., Lemaignan, S., Sisbot, E.A., Alami, R., Steinwender, J., Hamann, K. and Warneken, F.: Which one? Grounding the referent based on efficient human-robot interaction. In: *IEEE RO-MAN 2010*, pp. 570--575. (2010)
19. Reeves, B., and Nass, C.: *The media equation: How people treat computers, televisions, and new media like real people and places.* Cambridge University Press, New York. (1996)
20. Heider, F. and Simmel, M.: An experimental study of apparent behavior. *American Journal of Psychology* 57, pp. 243--249. (1944)
21. Weiss, A., Igelsböck, J., Tscheligi, M., Bauer, A., Kühnlenz, K., Wollherr, D., and Buss, M. Robots asking for directions – the willingness of passers-by to support robots. In *HRI '10: Proceedings of the 5th ACM/IEEE International Conference on Human Robot Interaction* (New York, NY, USA, 2010), pp. 23-30, ACM.
22. Barberis J.-M.: Indiquer son chemin au passant: rôle cognitif et discursif de orientation generale. In: Barberis, J.-M. (editor): *La Ville. Arts de Faire, Manieres de Dire, Langue et Praxis*, pp. 77--97. (1994)
23. Mondada, L.: Deixis spatiale, gestes de pointage et formes de coordination de l'action. In: Barberis, J.-M., Manes-Gallo, M.C. (Eds.), *Parcours dans la Ville. Les Descriptions D'itinéraires Piétons.* L'Harmattan, Paris, pp. 261--285. (2007)
24. Buchner, R. Förster, F., Mirnig, N., Weiss, A. and Tscheligi, M.: Deliverable 2.0@M6: Characteristics and properties of human-robot interaction in public environments. Tech. rep., <http://www.iuro-project.eu/>. (2010)
25. Byrt, T. How Good Is That Agreement? In *Epidemiology*, 1996, 7, 561. (1996).
26. Althusser, L.: *Ideology and Ideological State Apparatuses.* In: *Lenin and Philosophy and Other Essays.* Monthly Review Press. (1972)
27. Goffman, E.: *Behavior in Public Spaces. Notes on the Social Organization of Gatherings.* The Free Press, New York. (1963)
28. Shannon, C.E. and Weaver, W. , *The Mathematical Theory of Communication.*University of Illinois Press, 1963.
29. Huettnerrauch, H., Severinson-Eklundh, K., Green, A. and Topp, E. A.: Investigating spatial relationships in human-robot Interaction. In: *Proceedings of the IEEE/RSJ international conference on intelligent robots and systems (IROS 2006).* (2006)
30. Pacchierotti, E., Christensen, H. and Jensfelt, P.: Evaluation of passing distance for social robots. In: *Proceedings of IEEE Workshop on Robot and Human Interactive Communication (RO-MAN 2006).* (2006)
31. Förster, F., Weiss, A., and Tscheligi, M. Anthropomorphic design for an interactive urban robot – the right design approach? In *HRI '11: Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction* (New York, NY, USA, 2011), ACM.