Czech Technical University in Prague
Faculty of Electrical Engineering
Department of Computer Graphics and Interaction

Verbal Route Descriptions in Navigation of People with Visual Impairments

Doctoral Thesis

Ing. Jan Balata

Ph.D. programme: Electrical Engineering and Information Technology
Branch of study: Information Science and Computer Engineering
Supervisor: Doc. Ing. Zdeněk Míkovec, Ph.D.

Prague, August 2018

Thesis Supervisor:

Doc. Ing. Zdeněk Míkovec, Ph.D.
Department of Computer Graphics and Interaction
Faculty of Electrical Engineering
Czech Technical University in Prague
Karlovo náměstí 13
121 35
Czech Republic

Declaration

I hereby declare I have written this doctoral thesis independently and quoted all the sources of information used in accordance with methodological instructions on ethical principles for writing an academic thesis. Moreover, I state that this thesis has neither been submitted nor accepted for any other degree.

In Prague, August 2018

....
Ing. Jan Balata

List of Tables

List of Figures

Contents

Abstract

English This thesis focuses on the problem of navigation of people with visual impairments. There are many research directions in this field from sensory substitution, obstacle detection aids to different kind of navigation systems based on various input and output modalities. However, currently available navigation systems do not support navigation and orientation skills of people with visual impairments in the sense of supporting the creation of proper route knowledge.

This thesis aims to explore the role of verbal descriptions in navigation from the perspective of human-computer interaction. In a series of experiments, we provide insights into verbal description sharing habits and tele-assistance navigation strategies. Based on those insights, we design three navigation systems prototypes – a collaborative navigation system, a landmark-based navigation system, and a conversational navigation agent. This thesis contributes both on a theoretical level providing in-depth observations of the navigation process and also on a pragmatic level providing prototypes and guidelines for designers of the future navigation system for people with visual impairments.

Keywords Human-computer interaction, accessibility, navigation, visual impairment, verbal route description, spatial language, conversational user interface.

Česky Dizertační práce se zaměřuje na problém navigace osob se zrakovým postižením. Existuje mnoho výzkumných směrů od senzorické substituce, přes systémy vyhýbání se překážkám, po navigační systémy s různými vstupními a výstupními modalitami. Současně dostupné navigační systémy však nepodporují přirozené navigační a orientační schopnosti osob se zrakovým postižením ve smyslu podpory vytváření správného mentálního modelu trasy.

Tato práce se zaměřuje na výzkum role slovního propisu tras na navigaci z pohledu oboru styku člověka s počítačem (human-computer interaction). V sérii experimentů poskytujeme vhled do problematiky sdílení slovních popisů a strategií tele-asistivní navigace. Na základě těchto vhledů jsme navrhli tři prototypy navigačních systémů – kolaborativní navigaci, navigaci na základě významných bodů a konverzačního navigačního agenta. Přínos této práce je jak na teoretické úrovni poskytnutím detailních pozorování navigačního procesu ale i na pragmatické úrovni poskytnutím prototypů a doporučení pro návrháře budoucích navigačních systému pro osoby se zrakovým postižením.

Klíčová slova Interakce člověka a počítače, přístupnost, zrakové postižení, slovní popis trasy, prostorový jazyk, konverzační uživatelské rozhraní.

Acknowledgements

I would like to thank all participants of the studies for their patience and all the feedback they gave us which drove me in the right directions. I hope, they will benefit from the results of this thesis and that they will be more independent in their lives.

Next, I would like to thank my supervisor Doc. Ing. Zdeněk Míkovec, Ph.D. for his inspirational energy which motivated me during the hard times throughout the years of my studies, and our Department for the support we got for our research, and my colleagues who often assisted during the experiments.

Finally, I would like to thank my parents and grandparents for their support and nurturing my positive attitude toward science since early childhood. Last but not least, I would like to thank my wife for her infinite patience and loving support.

Chapter 1

Introduction

Abstract. *The mobility of a person is largely limited by visual impairment. This leads to a decrease in travel-related activities and less involvement in social life. People with visual impairment have different strategies regarding orientation in space, some of them are trained by orientation and mobility specialist, some of them are gained through time. While the construction of cognitive maps of the route is required for efficient navigation, without sight other strategies have to be employed. One of them is a verbal description of a route, which was showed to have a similar effect on the creation of cognitive maps as visual and direct experience.*

1.1 Motivation

Even with the huge leap of technologies in past years, there are still few commercially available navigation aids for people with visual impairments. In their daily travel-related activities, they end up relying on their trained skills, white cane, a smartphone with car navigation application, and the help of someone else (either directly on the street or via phone, see Subsection 3.2.1).

The ability to travel independently is required for the satisfactory level of quality of life and self-confidence. Visual impairment limits mainly a person's mobility and reduces travel-related activities [46]. Even though visually impaired people undergo special training of navigation and orientation skills, 30 % of them never leave home alone without a sighted guide [28], [112] and this number remains stable for the past twenty-five years despite the advances in information and geolocation technologies.

People with visual impairments experience a high level of stress whenever they try to travel independently [119], and only a fraction of blind people travel independently to unknown places [47]. Restricted mobility typically results in loss of leisure time activities (38 % of visually impaired people reported at least some interference with their leisure activities [101]), loss of career and work opportunities, and loss of social contacts [24]. The independent movement also plays an important role in the process of adjustment to blindness. According to the model defined by Tuttle and Tuttle [103], the adjustment process consists of seven phases: trauma, shock and denial, mourning and withdrawal, succumbing and depression, reassessment, and reaffirmation, coping and mobilization, self-acceptance, and self-esteem. Independent movement and the ability to navigate in known and unknown spaces is important for the sixth phase, coping and mobilization, where it can help to develop key competencies, such as the ability to work, to create social relations, or to act independently.

In general, the mobility of a person is mainly influenced by the efficiency of the wayfinding process. This process consists of three stages: planning of the route, execution of the planned route, and processing the information about the environment [62]. All parts of the wayfinding process can be supported by navigation aids, which assist the visually impaired pedestrian.

1.2 User Group Definition

The World Health Organization classifies the visual impairments according to severity on the range from mild or no visual impairment, thought moderate visual impairment and severe visual impairment (or low vision) to blindness (light perception) and total blindness (no light perception) [113]. Further, we may refer to people based on the onset of the disability as congenitally blind (born with the disability), early blind (acquired the disability in early childhood) and late blind (in higher age, i.e., after puberty).

There are 253 million people with visual impairments. Thirty-six million of them are blind, and 217 million have moderate to severe vision impairment. Around 81 % of people who are blind or have moderate or severe vision impairment are aged over 50 years [19].

First, it is important to get insights into the different aspects playing a role in the orientation and mobility skills of people with visual impairment. The key resource is research conducted in past decades in the field of cognitive psychology.

1.2.1 Orientation and Navigation

There are three levels of spatial knowledge: landmark knowledge which is the simplest level of the spatial knowledge, holds information about the existence of important places and objects and ability to recognize them [52]; survey knowledge which captures relations between landmarks represented by angles or distances [111]; and route knowledge which consists of series of procedural descriptions about origin, destination and landmarks along the route [42].

In large spaces where body movement is necessary, visually impaired people use different cognitive strategies from those used by sighted people for navigation and orientation. These strategies are mostly based on egocentric frames [68], [69]. Typically, they have to memorize a large amount of information [102] in the form of sequential representation [69] based on routes. Route knowledge has to be acquired on a declarative level [47] which means, that the route knowledge can also be used when the person is not on the route itself; on the other hand, procedural level route knowledge denotes the route knowledge which is in a form of a fixed automated reaction to stimuli on the route.

Fortunately, it seems that visually impaired people acquire superior serial memory skills. A study by Raz et al. [84] discovered that congenitally blind people are better than sighted people in both item memory and serial memory and that their serial memory skills are outstanding, especially for long sequences. In a study by Bradley and Dunlop [20], it was revealed that in a situation of pre-recorded verbal navigation, the blind navigator navigated the blind traveler significantly faster than a sighted navigator.

1.2.2 Cognitive Mapping

According to Downs [35]: "Cognitive mapping is a construct which encompasses those cognitive processes which enable people to acquire, code, store, recall, manipulate information about the nature of their spatial environment, and is an essential component in the adaptive process of spatial decision making. Similarly, a cognitive map is an abstraction, which refers to a cross-section, at one point in time, of the environment as people believe to be."

There is clear distinction among people who have visual memory (e.g., lost their sight during the life) and people who are congenitally blind. The late blind people have some visuospatial memories and tend to navigate similarly to sighted people for a certain period of time. On the other hand, the congenitally blind people without visuospatial memories had to developed different strategies to cope with navigation.

Another important aspect in navigation and orientation of visually impaired people is the difference between near and far space in which they perform a particular task. The near space – haptic space – relates to small-scale or manipulatory space i.e., areas that can be explored without changing a position of the body. The far space – locomotor space – relates to medium- or large-scale space e.g., areas in which locomotion is required for exploration [105].

1.2.3 Coding Strategies

Various coding strategies are important for exploring space around a person in the nearest vicinity. They are also employed in exploring tactile maps, which are proved to be useful in training of new routes and exploring unknown environments [38] sometimes enhanced with multimodal (vibration or sound) information [21].

There are two options for coding the location of an object in haptic space: either by reference to our own body and movements or some external network. People with little or none visually experience (usually congenitally blind and early blind people) tend to code spatial relations by reference to their own body. On the other hand, for sighted people it is more natural to code the position to an object relative to other objects (external network) [105].

Coding strategies based on the reference to a body are more reliable in haptic space and are equivalently efficient as strategies used by sighted people [43].

1.2.4 Wayfinding and Cognitive Maps in Locomotor Space

From the various methods used to investigate cognitive maps of blind people, the most common is a direct reproduction of the route.

It was found that blind children were less accurate in reproducing familiar environment (school campus) than sighted ones, however, there were some individual differences when blind participants were more accurate than sighted [25]. It was later proven that the congenitally blind participants had a tendency to linearize curved paths, that the maps were segmented and the features of the routes were more accurate on familiar routes.

Some of the studies of locomotor space were focused on the problem of exploration of the unknown space [44], [54]. They have identified several strategies observing the behavior of the blind participants:

- **Perimeter.** Explored the boundaries of an area to identify the area's shape, size and key features around its perimeter, by walking along the edge of the layout.
- **Grid.** Investigated the internal elements of an area to learn their spatial relationships, by taking straight-line paths from one side of the layout to the other.
- **Object to object.** Moving repeatedly from one object to another, or feeling the relationship between objects using hand or cane.
- **Perimeter to object.** Moving repeatedly between an object and the perimeter.
- **Home base to object.** Moving repeatedly between the home base (origin point for exploration) and all the others in turn.
- **Cyclic.** Each of the four objects visited in turn, and then returning to the first object.
- **Back-and-forth.** Moving repeatedly between two objects.

Perimeter and grid strategies used in isolation gave good knowledge of object location, however, in a test of integrated spatial knowledge, these strategies had worse results. The most successful participants tended to use an object to object, perimeter to object or home-base to object strategies, and also often used a wider range of different strategies. Cyclic patterns were used mainly by early blind participants (individuals who were born blind or lost sight before the age of 3 when visuospatial concepts were not fully developed), whereas late blind and blindfolded sighted participants tended to use the back-and-forth strategy, which was associated with good performance. Conversely, cyclic exploration was associated with a poorer performance.

Generally, the experiments with blind participants show that the lack of visual experience does not prevent them from the acquisition of spatial representation and creation of cognitive maps. Also, it was proven that totally blind participants perform at the level of sighted participants on spatial tasks. Spatially relevant information is available through senses other than vision, like through hearing, touch and movement, and this information can form the basis for spatial coding. An important fact is that the processing of spatial information by congenitally blind people is not necessarily less efficient than by sighted people [105].

1.2.5 Verbal Descriptions

Denis and Zimmer [33] showed that verbal description of a route can be adequately converted into mental representations and that it creates similar results to the visual experience. Moreover, the mental representations created in both ways show similar metric properties [31].

Similarly, Striuksma [98] discovered that even congenitally blind people can perform tasks traditionally associated with spatial imagery, showed that the connection between different modalities is used to form an abstract mental representation, and suggested that spatial images are supramodal. On the other hand, Loomis et al. [64] later argues that spatial images are amodal and that their functional characteristics are not tied to a source modality, i.e., their functional properties are equivalent. Verbal descriptions of a route are tightly connected to spatial language, which is defined as "a means of representing objects and locations through verbal description with respect to multiple coordinate systems or frames of reference." [18]. Spatial language has an important role in wayfinding process: it is used first, to form a mental representation of space; second, to plan a route; third, to formulate an action to reach the goal [32] (which corresponds to stages of wayfinding process [78] in context of the visually impaired population). Namely, Denis [32] observed the usage of spatial language in the context of referring to action, referring to a landmark, referring to an action together with a landmark, description of a landmark, and commentaries.

Similarly, Golledge et al. [47] claim that one of the basic strategies for independent and efficient navigation is navigation from landmark to landmark. Landmarks are defined as conceptually and perceptually distinct locations, build or naturally occurring (e.g., a church on a square, or specific smell of nearby bakery). Landmark may not be only one feature but also a combination of several features creates a landmark [47]. Not surprisingly, landmarks are also the most prevalent information in the verbal description of routes used by humans and were also proved to increase the confidence and trust of the traveler when incorporated to the route description [67]. The choice of landmarks is subject to individual capabilities of the traveler – in case of visually impaired travelers, a suitable landmark can be a leading line (natural or artificial), corner of a building with a specific shape, a location of the busy street, or specific material or slope of a sidewalk. Further, Ross et al. [88] have shown that inclusion of landmarks within pedestrian route itinerary increased user confidence and reduced navigation errors. Findings of Ross et al. [88] also extend to voice-only navigation [85], where participants preferred inclusion of landmarks. There are also some experimental designs of navigation systems, which rely primarily on landmarks to navigate users from origin A to destination B, e.g., Millonig and Schechtner [70].

For the successful learning of new routes and new environment, the survey knowledge has to be formed [94]. Noordzij [75] showed that verbal description positively supports the creation of survey knowledge.

1.3 Research Problem and Research Questions

As we can see, there are many differences between orientation and mobility of people with visual impairment and the general population. The approach we take in this work should follow and respect their navigation skills and strategies. Thus, we choose a User Centered Design (ISO 9241-210:2010) [58] methodology as a center point of the work to be able to reflect the specific needs of our user group. Although a lot of work was done in this field, not many influenced development of commercially available navigation aids (see Related Work in Chapter 6).

> **Research problem:** We identified a problem in how the current navigation systems support the specific needs of visually impaired pedestrians. Particularly in the way that they do not provide a verbal description that would support them in gaining sufficient route and survey knowledge or the new route.

This thesis aims the provide answers to following research questions which are tightly connected to our approach of solving the research problem:

- **RQ1** How do people with visual impairments share spatial knowledge in a form of verbal route description?

- **RQ2** How much and how well do people with visual impairments remember their regular routes?
- **RQ3** How do people with visual impairments perform while sharing the verbal route description over the phone?
- **RQ4** How can be tele-assistance navigation for people with visual impairments improved in general?
- **RQ5** How will people with visual impairments perform with automatically generated verbal description during navigation?

1.4 Organization of the Thesis

The thesis is organized into seven chapters. First three chapters provide a reader with theoretical framework and setting in which the following four chapters introduce conducted research providing answers to defined research questions.

- Chapter 1 describes the motivation and general related work.
- Chapter 2 summarizes the methodology and methods used in this thesis.
- Chapter 3 reports on the user studies focused on sharing of route verbal description in the community of people with visual impairments.
- Chapter 4 reports on the experiment where visually impaired people navigated each other over the phone (tele-assistance).
- Chapter 5 presents a concept of collaborative navigation for visually impaired people.
- Chapter 6 builds upon the third and fourth chapter. It presents an approach to automatically generate verbal route description with similar properties to human-made ones.
- Chapter 7 describes the design and prototyping process of a conversational agent using outcomes of third and sixth chapter along with a description of methods to use to enable global and local personalization for the users.
- Chapter 8 concludes the thesis.

Research questions RQ1, RQ2 are answered in Chapter 3, RQ3, RQ4 in Chapter 4 and RQ5 in Chapter 6. Chapters 5 and 7 introduce concepts and prototypes implied by the research questions.

Chapter 2

Methodology and Methods

Abstract. *In this chapter, we summarize the research methodologies and methods used to provide answers to the research questions defined in the previous chapter. Besides the focus on answering the research questions, our research also focuses on providing practically applicable outcomes to directly benefit the user group.*

2.1 Methodology

This thesis is placed within the pragmatic research paradigm [66]. Both qualitative, quantitative, and mixed methods of data collection, analysis and reporting are used to best fit the research problem. The focus is on understanding and answering the research questions, solving emerging problems, and delivering practically applicable results. We employ User-Centered Design methodology, and we strongly focus our aim around the users.

We set this thesis around two key frameworks: First, Ability-based design [118] mainly the principle of Ability – leveraging all the abilities of a user – and the principle of Commodity – creating a low-cost solution using readily available hardware (and software) solution. We focus on building systems, which support natural orientation skills and strategies of visually impaired and use current technologies without the need for buying or using additional hardware.

Second, the quantitative difference theory [40] where we assume the normal range of spatio-cognitive behavior of visually impaired people, even though some [68], [99] argue there is a thin border between deficiency theory and inefficiency theory [96] depending on the particular task and its context.

2.2 Methods

We use different types of research methods to answer the research questions. We often use a combination of the methods, e.g., experimental accompanied by a correlational method. For research questions RQ1, RQ3, RQ4, and RQ5 we use observational methods. For RQ1, RQ2, RQ5 correlational method. For RQ5 we use experimental methods.

RQ1 (How do people with visual impairments share spatial knowledge in a form of verbal route description?) is approached first by semi-structured interview from which we formulated hypotheses, which were then assessed using a survey. We selected an interview instead of a focus group as we wanted to get deeper into the problem field. Furthermore, interview enabled us to invest more time into individual participants and obtain deeper insights.

RQ2 (How much and how well do people with visual impairments remember their regular routes?) is approached by means of a quantitative structured interview. We selected this method as we aimed to collect highly complicated structured data, for which a in-person presence of a researcher was required (given the user group and accessibility issues connected with for example remote survey) – we asked the participants to name routes they take every week and then let them describe each of the routes in as much detail as possible.

RQ3 (How do people with visual impairments perform while sharing the verbal route description over the phone?) and RQ4 (How can be tele-assistance navigation for people with visual impairments improved in general?) are handled together using a field study method to observe the behavior of the participants in the natural environment and reach higher ecological validity of the research results.

RQ5 (How will people with visual impairments perform with automatically generated verbal description during navigation?) was approached in three separate experiments. In the first experiment, a mixed method was used: an empirical study when we systematically changed an independent variable combined with a questionnaire (Likert scale). Next, a field study was performed, and qualitative data were collected using contextual inquiry when interacting with a prototype. Last, a diary study method was applied to observe a long term interaction with a prototype.

Chapter 3

Verbal Route Description Sharing

Abstract. *A navigation system for visually impaired users may be much more efficient if it is based on collaboration among visually impaired persons and on utilizing distributed knowledge about the environment in which the navigation task takes place. To design a new system of this kind, it is necessary to make a study of communication among visually impaired users while navigating in a given environment and on their regularly walked routes. A qualitative study was conducted to gain insight into the issue of communication among visually impaired persons while they are navigating in an unknown environment, and our hypotheses were validated by a quantitative study with a sample of 54 visually impaired respondents. A qualitative study was conducted with 20 visually impaired participants aimed at investigating regularly walked routes used by visually impaired persons. The results show that most visually impaired users already collaborate on navigation, and consider an environment description from other visually impaired persons to be adequate for safe and efficient navigation. The results also suggest that the proposed collaborative navigation system is based on the natural behavior of visually impaired persons. In addition, it has been shown that a network of regularly walked routes can greatly expand the urban area in which visually impaired persons are able to navigate safely and efficiently.*

Note. *This chapter is based on published journal paper Collaborative Navigation of Visually Impaired [8].*

3.1 Introduction

The main difference between navigation systems specially designed for sighted users and visually impaired users lies in the level of detail of the environment description, and in the representation of the instructions.

Sighted users and drivers are bound to streets and roads, and the details contained in a description of these elements are sufficient for them. However, visually impaired users make use of different navigation features in the environment, and the description provided by a navigation system specially designed for sighted persons is not adequate for them [108].

Several navigation applications on smartphones are nowadays used by visually impaired persons, although they were originally designed for sighted pedestrians or drivers. The most common is the Nokia Maps application, pre-installed on Symbian[1] smartphones, which is widely used by visually impaired users because of their hardware keyboard and their good screenreader[2]. Other options are built-in navigation applications either in Android[3] smartphones (version 4.0 and above) with Explore-by-Touch support for eyes-free interaction, or in iOS[4] devices with VoiceOver[5].

Situations in which orientation is lost are mentally demanding, especially for visually impaired users, and it is necessary to analyze ways to help them by means of navigational aids. Some partial solutions to this problem are available in the form of:

- navigation call centers (e.g., Navigation Center of Czech Blind United organization [30] where orientation and mobility specialists and trained operators directly navigate visually impaired persons),
- voice-enabled navigation applications for smart phones (i.e., Loadstone GPS[6] or BlindSquare[7]),
- custom hardware or wearable computer aids [56], [83], [123].

The last of these three solutions can involve a wide range of sensors, e.g., cameras, headsets for communication, GPS, or even Microsoft Kinect, to analyze images and provide information for visually impaired users about their surroundings.

Although all these navigation systems are designed for visually impaired persons, none of them is ideal. Important considerations are that many custom hardware or wearable computer aids are in the prototype phase of development, and only a small proportion of all aids can resolve situations in which a visually impaired user gets lost.

3.1.1 Use Case

Let us imagine that Alice is a visually impaired person who has decided to visit a new art exhibition for the visually impaired in the city center on a Saturday evening (see Figure 3.1). She knows the address, but she does not know the destination area or the route well. She does not have a computer, and she uses a smartphone equipped with a GPS module and the installed screen reader for all her communication (calls, messages, emails, web). If she feels unsure about the route, or if she gets lost, she might be able to use several available navigation methods:

- ask strangers passing by,
- use the voice-enabled navigation application on her smartphone,
- call the navigation center for the visually impaired and get instructed remotely,
- get in touch with an orientation and mobility specialist,

---

[1] Symbian OS – https://en.wikipedia.org/wiki/Symbian

[2] Mobile Speak – http://codefactoryglobal.com/app-store/mobile-speak/

[3] Android OS – https://www.android.com/

[4] iOS – https://www.apple.com/lae/ios/ios-11/

[5] VoiceOver – https://www.apple.com/lae/accessibility/iphone/vision/

[6] Loadstone GPS – https://www.loadstone-gps.com/

[7] BlindSquare – http://www.blindsquare.com/

Figure 3.1: Alice is a visually impaired person who has decided to visit a new art exhibition for the visually impaired in the city center on a Saturday evening but does not know the destination or the route well.

- call visually impaired friend who may know the destination area and who will instruct her remotely,
- call an unknown visually impaired person with useful knowledge of the desired destination are, who walks past the gallery (where art exhibition takes part, see Use Case) every day on the route from his/her home to a nearby public transport stop to get to work (there may be more than one visually impaired person able to help in this way).

3.1.2 Problem Description

A common problem for all the existing navigation solutions is the lack of an environment description especially created for visually impaired users, focusing on specific navigation points (e.g., leading lines) and orientation cues (e.g., sounds, slopes), i.e., landmarks [46], [105], [108], which visually impaired users need for safe and efficient navigation. The mental representation of the route can be formed from direct experience or a verbal description [32]. The verbal description of this kind is hard to obtain, as only trained orientation and mobility specialist can provide it in a fully satisfactory form and it requires good knowledge of the area and necessity to perform on-site reconnaissance. Unfortunately, there is typically only a very limited number of such specialist.

A useful description can also be provided by visually impaired persons who are familiar with a particular area. Typical landmarks provided by visually impaired people based on their own experience and observation are:

- leading lines formed by edges of the pavement, handrails, corners of buildings, the better side of the street to travel on, etc.,
- sounds from traffic, construction work, water, width of the street from the echo, etc.,
- smells of various types, e.g., bakeries, drugstores, sewers, etc.,
- the direction of heat from the sun at certain times of day,

- approximate distances, changes in elevation, type of ground material.

For the names of streets, the order of turns on the route, navigation through parks and estates, the location of a specific house in the street, etc. visually impaired people need a description from an orientation and mobility specialists.

People with visual impairments remember the verbal description of frequent routes provided by orientation and mobility specialist and can pass these descriptions to other people. There are much more visually impaired people with specific route knowledge than orientation and mobility specialists available. The problem is that the knowledge is fragmented to a number of rather short routes.

An idea for overcoming these problems is to support and facilitate direct simultaneous help between visually impaired people, and sharing of their knowledge about certain places (navigation points and orientation cues).

In this chapter, we present a set of experiments focused on exploring how people with visual impairments share a verbal description of their regular routes and how much of the regular routes they remember and can efficiently share.

3.2 Related Work

The related work sections summarizes part of related work focused on collaboration in navigation of visually impaired.

3.2.1 Collaboration in Navigation

The collaboration with and among visually impaired people has been explored and evaluated in several system and prototypes. One of the most used was VizWiz [17], an object recognition system where sighted volunteers and crowdworkers [87] helped visually impaired people recognize the object. Similarly, Be My Eyes[8], which connects visually impaired people with sighted volunteers, is commonly used in situations where image recognition or OCR does not work [6].

More into the area of navigation, Hara et al. [50] used sighted volunteers to describe public transport stations location or entrances to buildings [122] to visually impaired people, or barriers on sidewalks and crosswalks [51] to hint wheelchair users about accessibility.

The collaboration also exists among visually impaired people themselves. Visually impaired people create a detailed mental map of their regular routes [79]. The regular routes are those learned by orientation and mobility specialist, self-learned or shared by friends or acquaintances [16], [82]. Landmarks are naturally occurring or build features of the environment that stand-out in the vicinity. They are often used when visually impaired people describe their regular routes to another person.

There are clear reasons for the statement that being navigated by another visually impaired person is perceived better. Williams et al. [116] observed the situation when sighted persons helped visually impaired person with navigation and strategies, language, and understanding differed largely. This is also supported by Bradley and Dunlop [20], who showed that visually

---

[8] Be My Eyes – https://www.bemyeyes.com/

Participant	**P1**	**P2**	**P3**	**P4**	**P5**
Age	57	63	30	25	34
Gender	M	F	M	M	M
Category	5	5	5	5	5
Onset	L	L	C	C	L

Table 3.1: List of participants, their age, gender, the category of impairment (5 – no light perception [113]), and the onset of impairment (C – congenitally blind, L – late blind).

impaired people navigate with less workload with a verbal description of route formed by another visually impaired person than by sighted person.

Based on the difference between the mental representation of the routes between sighted and blind people we draw an inspiration for navigation based on collaboration among visually impaired people themselves by means of sharing their regular routes on the fly via a phone call, i.e., tele-assistance navigation.

3.3 Ways of verbal description sharing

First, we aim to gain in-depth insights into the way visually impaired people share their regular routes, how they are willing to help others and how much information they are able to share in sufficient quality.

3.3.1 Qualitative study

The goal of the qualitative user study was to obtain in-depth information about participants favorite means of communication with other people, their openness to communicating with strangers, their willingness to help them with navigation, and the privacy problem of the location sharing and storing.

Participants

Five visually impaired participants (1 female, 4 males) were recruited via an e-mail leaflet sent to a group of Czech Blind United organization [30] clients (see Table 3.1). The participants in the user study were aged from 25 to 64 years ($mean$ = 41.80, SD = 17.05). All of the participants had Category 5 vision impairment (no light perception) [113]. Two participants were congenitally blind, and three participants were late blind.

Procedure

We conducted five semi-structured interviews (all of them were conducted in person). The interviews lasted 60 minutes and were performed in a usability lab dedicated for execution of user studies. We asked the participants about the community of visually impaired persons, communication in navigation, experiences with locations services, privacy issues of location sharing, and behavior in situations of loss of orientation. After the interview, the participants were debriefed and received their payment.

No.	**Factoid**	**Occurrence**
1	Participant has experience of navigating other visually impaired persons, and has been navigated by other visually impaired persons by mobile phone, ICQ, etc.	5
2	Navigational instructions given by a visually impaired person about a place that he/she knows are better than instructions from a sighted person.	4
3	Asking strangers in the street for directions is a natural part of the navigation process.	4
4	Participants have concerns about location sharing.	3

Table 3.2: Key factoids extracted from the semi-structured interviews.

Findings

We extracted 4 key factoids from the semi-structured interviews (see Table 3.2). Further we report overall findings.

Community of visually impaired persons. More than half of the participants meet with their friends regularly, and the majority of the participants know where their family members and good friends live. Some of the participants regularly visit places connected with their hobbies. For example, P1 stated: “I used to ski, but it was a long time ago. Later I went to yoga class. Now I visit rehabilitation exercises every Tuesday.”

Communication in navigation. Majority of the participants had experience with navigation other visually impaired persons remotely and had themselves been navigated by another visually impaired person or by a sighted passerby (No. 1). Some of the participants requested navigation by a phone call and some by an e-mail or SMS. For example P1 said: *“We advise over the phone each other. For example, I know that this person knows the route to a place where I want to go. So I ask him to describe to me the route because he knows all the important landmarks.”* Most of the participant agreed that navigation instructions given by a visually impaired person are better than the instructions given by sighted persons (No. 2).

Experience with location services. More than half of the participants had some experience with GPS navigation applications in smartphones. Some of them use the navigation application regularly (either on known and unknown routes), despite the low GPS accuracy [71] in city centers.

Privacy issues of location sharing. More than half of the participants expressed concerns about sharing their location even for the purposes of helping other visually impaired persons (No. 4). For example, P1 stated: *“I don’t like the idea that someone would know where I go regularly.”* or *“If there was an agreement about the data usage, I wouldn’t mind if some other visually impaired person contacted me with a request for navigation on the route I know.”*

Behavior in situation of going astray. Most of the participants consider asking a passerby for directions to be a natural part of the navigational process (No. 3). However, there was a case where a participant (P1) hesitated to ask a passerby: *"When I get lost, I have to ask a passerby. However, I try to postpone it as far as it goes. First I try to back-track to find the right way."* and *"Once, when there was a big snowfall, I shouted outside for a help. I didn't try the phone, because I didn't know where I was at all."*

One of the participants also mentioned that there is a problem with finding someone suitable to ask for help (some passerby do not want to help him/her, or do not speak Czech) and for this reason he/she does not like to ask people in the street.

Hypotheses

From the key factoids described in Table 3.2 we define following hypotheses, which we evaluated in a survey 3.3.2 described in the next subsection within this chapter.

- **H1** A visually impaired person can navigate another visually impaired user via mobile phone.
- **H2** A visually impaired person will prefer navigation instructions provided by another visually impaired person to instructions from a sighted person.
- **H3** A visually impaired person will not hesitate to ask strangers in the street in order to get reliable directions.
- **H4** A visually impaired person will allow information to be collected about his/her location in order to help in navigating other visually impaired people.

3.3.2 Quantitative Survey

The quantitative study was conducted on the basis of a questionnaire compiled from the findings obtained in the qualitative studies. The goal was to validate the hypotheses formulated on the basis of the qualitative user study.

Participants

Fifty-four visually impaired respondents (18 females, 36 males) were recruited via an e-mail leaflet sent to a group of Czech Blind United organization [30] clients. The respondents were aged from 20 to 80 years ($mean = 47.06$, $SD = 15.89$). The respondents had Category 4 vision impairment (light perception) and Category 5 vision impairment (no light perception) [113]. Seventeen respondents were congenitally blind, and thirty-seven respondents were late blind.

Procedure

We conducted a survey using questionnaire with 12 closed-ended questions. The questions were administered via e-mail. The respondents filled-in the responses in a reply to the original e-mail. The general topics of the questionnaire were demographic information, behavior while obtaining navigation information from passerby in the street, preferences in selecting contacts when seeking for navigation instruction, navigating friends and family members via phone, email, instant messaging, etc., requesting help with navigation from friends or family members, willingness to help an unknown visually impaired person with navigation, and privacy problem of location sharing.

Findings

Fifty-four e-mails returned with a filled-in questionnaire. Due to the nature of the impairment of the participants' not all answers were filled-in correctly. The exact number of respondents taken into account in each finding is reported in the results, only correctly filled-in are taken into account.

Do you ask passersby on the street for navigation instructions (n = 47)? The percentage of the respondents who selected particular answer follows: 49 % – often, I find it a natural part of the navigation process; 47 % – I do not like asking passerby, first i try to find the way myself; 4 % – I do not ask passerby.

How often do you call your friends or family members for help with navigation (n = 49)? The percentage of the respondents who selected particular answer follows: 49 % – never; 41 % – several times a year; 8 % – several times a month; 2 % – daily.

How often do you help your friends or family members with navigation (via phone call, ICQ, e-mail, etc.) (n = 49)? The percentage of the respondents who selected particular answer follows: 53 % – several times a year; 33 % – never; 10 % – several times a month; 2 % – several times a week; 2 % – daily.

If an unknown visually impaired person calls you for help with navigation at the location you walk daily, will you help him/her (n = 49)? The percentage of the respondents who selected particular answer follows: 82% – yes; 12 % – probably yes; 4 % – I do not know; 2 % – no.

Would you mind automated recording (e.g., using GPS) of your regularly travelled routes in order to help other visually impaired in navigation (n = 50)? The percentage of the respondents who selected particular answer follows: 68 % – do not mind; 12 % – rather do not mind; 8 % – do not know; 8 % – rather mind; 4 % – mind.

Do you use electronic navigation (i.e., smartphone application) (n = 50)? The percentage of the respondents who selected particular answer follows: 68 % – do not use; 18 % – several times a month; 12 % – several times a year; 2 % – daily.

Preferences in selection of the person who will help the visually impaired person to navigate in an unknown place (n = 31). The score calculated from the ordering of the following four variants: 3.52 – sighted person with experience of navigating visually impaired persons; 3.26 – blind person who knows the place of navigation; 2.19 – sighted person with no

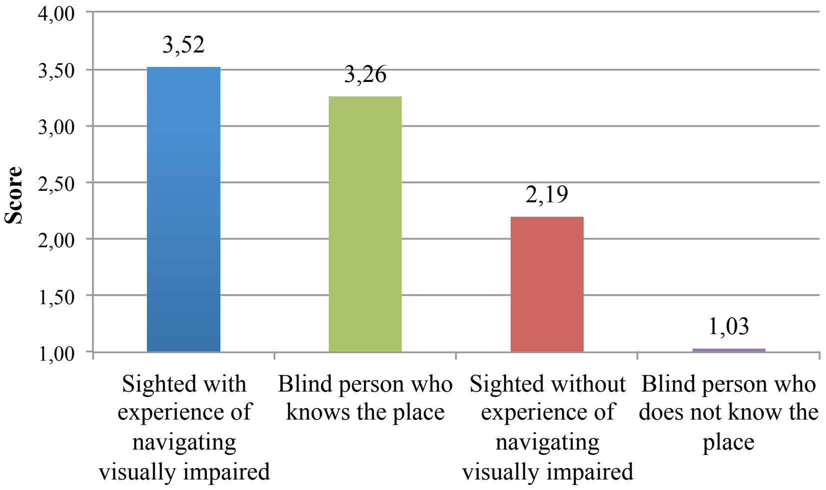


Figure 3.2: The preference for selecting a contact for help in navigation (n = 31).

experience of navigating visually impaired persons; 1.03 – blind person who does not know the place of navigation (minimum = 1, maximum = 4).

3.3.3 Discussion

It seems that hypothesis H1 has been proven, as the visually impaired users reported that they can navigate or be navigated by other visually impaired users. The results show that visually impaired users already collaborate on navigating in small groups of users – family or groups of friends. We believe that appropriate tools and aids could further enhance the collaboration of these groups. These tools and aids should help to connect visually impaired users who do not know each other but can collaborate in navigating.

Hypothesis H2 is valid if we only take into consideration untrained sighted persons. H2 is not valid for expert sighted persons. Figure 3.2 shows that visually impaired people rate the quality of the description provided by a sighted expert more highly than the description provided by a visually impaired person who knows the area. Respondents seem to expect that the description provided by a sighted expert will be more customized to their needs, as the expert has more information sources available.

The validity of hypothesis H3 is not proved. 51 % of the participants hesitate to ask strangers for instructions for navigation. However, the size of this group of participants can be attributed to a fact mentioned by one of the participants in the qualitative study – the problem of finding a suitable person who is willing to help. We believe that this problem can be solved by implementing our future system, where all of the users proposed by the system to provide help will have valuable information for navigation and will be willing to help.

In order to select accurately a suitable person to provide the right information for a visually impaired person who is lost, the location of all users needs to be stored and analyzed. The results shows that 68 % of all respondents have no problem with location sharing and location storing. If we add in those participants who do not mind only to some extent (12 %), a total of 80 % of the participants provide support for hypothesis H4.

ID	Age	Gender	Category	Onset	ID	Age	Gender	Category	Onset
P01	36	M	5	L	*P11*	46	M	5	L
P02	64	M	5	L	*P12*	47	M	5	C
P03	65	F	5	L	*P13*	24	F	5	C
P04	37	F	4	L	*P14*	35	M	4	L
P05	31	M	5	C	*P15*	27	M	4	C
P06	37	F	4	L	*P16*	56	F	4	C
P07	59	M	5	L	*P17*	26	F	4	C
P08	47	F	5	C	*P18*	72	M	4	L
P09	30	M	5	L	*P19*	41	F	5	L
P10	46	F	4	L	*P20*	20	F	5	C

Table 3.3: List of participants, their age, gender, the category of impairment (4 – light perception, 5 – no light perception [113]), and the onset of impairment (C – congenitally blind, L – late blind).

3.4 Length of the Regular Routes

A regular route is typically the route to work, to the local shopping area, to the nearest public transport stop, or a route that a visually impaired person uses for taking a walk. In our definition, it is a route which is walked at least once a week. Usually, the person knows every detail of the regular route, e.g., the optimal leading lines, cracks in the pavement, smell and sounds (echoes), strategies for crossing a street, etc.

These routes can be collected by tracking the person or can be collected via an interview and plotted into a geographical database. These two methods can be used to produce a map showing the density of regular routes of multiple people, which can be then rendered as a heat map. Areas with a greater density of regular routes have a greater number of visually impaired people regularly walking through that area so that higher-quality verbal descriptions of the routes are available for navigation.

3.4.1 Quantitative Structured Interview

The aim of the quantitative user study was to identify regular routes walked by the visually impaired participants, measure the approximate length of the regular routes, estimate the contribution of the regular routes walked by the visually impaired persons to the current network of public transport.

Participants

Twenty visually impaired participants (10 females, 10 males) were recruited via an e-mail leaflet sent to a group of Czech Blind United organization [30] clients (see Table 3.3). The participants were aged from 20 to 72 years ($mean = 42.30$, $SD = 14.81$). Eight participants had Category 4 vision impairment(light perception) and twelve participants had Category 5 vision impairment (no light perception) [113]. Nine participants were congenitally blind, and eleven respondents were late blind. Two participants used a guide dog. Sixteen participants lived in Prague, and four participants lived in other large cities. All of the respondents were native Czech speakers.

Procedure

We conducted a 20 semi-structured interviews (all of them were conducted in person). The interviews lasted for approximately 40 minutes and were performed in a usability lab dedicated for execution of user studies. We asked the participants to describe their regular routes walked at least once a week in as much detail as possible, as if they were describing the route to another visually impaired person. After the interview, the participants were debriefed and received their payment.

Results

After the interviews we reconstructed all the collected regular routes from the description into a map, we measured the length of the routes for each participant and created a heat-map (see Figure 3.4).

The twenty visually impaired participants who were interviewed walk a total of 93.33 km of regular routes, and were able to describe the environment of these regular routes. This sum represents the length of unique regular routes for all participants (e.g., the route to the shop twice a week counts as one unique regular route). The mean length of the routes per participant is 4.67 km. The participants in the study visit 4.25 destinations, on an average (e.g., restaurants, workplaces, schools, hobbies, shops, etc.). The nearest public transport stops from home are not counted as destinations if they lie on the way to other places.

Regarding the level of detail of the verbal descriptions, the participants described common landmarks such as slopes, leading lines, auditory and olfactory landmarks, names of streets, cracks in sidewalks, walls, the position of buildings, or directions. All descriptions were in such detail we were able to reconstruct all the routes and draw them into a map. However, quantitative analysis of landmark usage was not performed.

The data suggests that there is no effect of congenital blindness and late blindness on the length of the regularly walked routes. There is a difference in the mean length of the routes for the late blind group, whose members walk 1 709 m further than the congenitally blind group. We think this may be partly due to inertia from the time when the participants were sighted, and when they had greater mobility and traveled more around the city. On the other hand, the group of congenitally blind persons tends to optimize from birth and to use the best (safest) route. They also tended to use public transport as much as possible. However, this assumption needs to be further investigated in a study of the different cognitive mapping strategies employed by congenitally blind and late blind persons.

Figure 3.3 shows the mean length of the regular routes for each participant. As expected, there are differences among the participants. Some of the participants walk long routes to the nearest public transport, or just for recreational walking. On the other hand, other participants avoid walking and use public transport as much as possible or use a guide for unknown or complicated places. It seems, there is no effect of age on the length of regular routes.

Figure 3.4 shows a heat map, which maps a number of routes in a certain place from blue to red values. Red areas indicate places where there are larger numbers of regularly walked routes involving one participant or more.

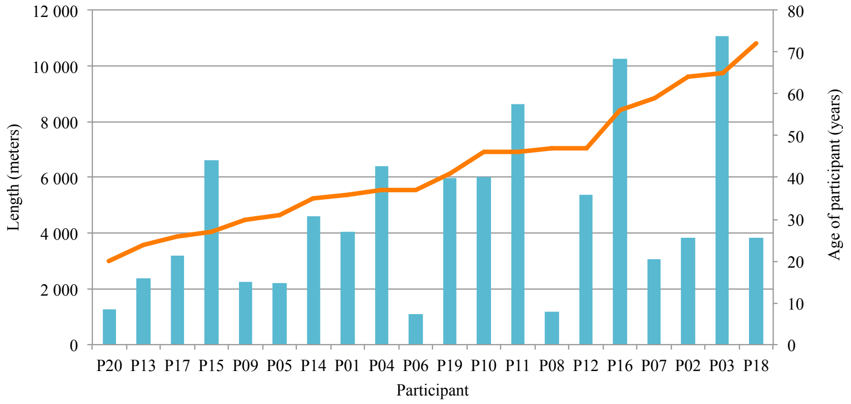


Figure 3.3: The relation between the age of the participant (orange line) and the length of the regularly walked routes (blue column) (n = 20). Participants are ordered by age in an ascending manner.

3.4.2 Discussion

Many of the regular routes of the participants start or end at public transport stops, which are usually easily accessible for visually impaired people. The total length of the public transport network in Prague is 1029.8 km (subway 59.4 km, tram 142.4 km, bus 828 km) [36]. The regular walking routes taken by the 20 visually impaired persons who participated in the study have a combined total length of 93.3 km. These walking routes extend the current public transport network by 9 % and create access to areas around public transport stops.

To cover the area accessible nowadays by public transport by regular routes, we would require approximately 200 active users for Prague (with none of them sharing the same route). The Czech Blind United organization [30] provides an estimated number of 100 000 people with visual impairments living in the whole Czech Republic. Considering that one-tenth of the population lives in Prague, there will be at least one thousand visually impaired persons in the city.

3.5 Conclusion

The results from the first study have revealed that most of the hypotheses about communication among visually impaired people during navigation were correct and provided important insights into the problems (see Section 3.2.1).

The data collected from the second study on regularly walked routes showed that visually impaired persons remember quite long route descriptions (in cognitive maps) and that they are able to describe them to other people with all important navigation points and orientation cues (see Section 3.4).

In summary, our studies have shown that both communication in navigation and regular routes of visually impaired people exist. They communicate about navigation on a regular basis and the regular routes have potential cover large areas of the city. These were the essential conditions for attempting to set up a navigation system based on collaboration among visually impaired persons. Both essential conditions have been fulfilled, and it has been shown that successful operation of the system is feasible.

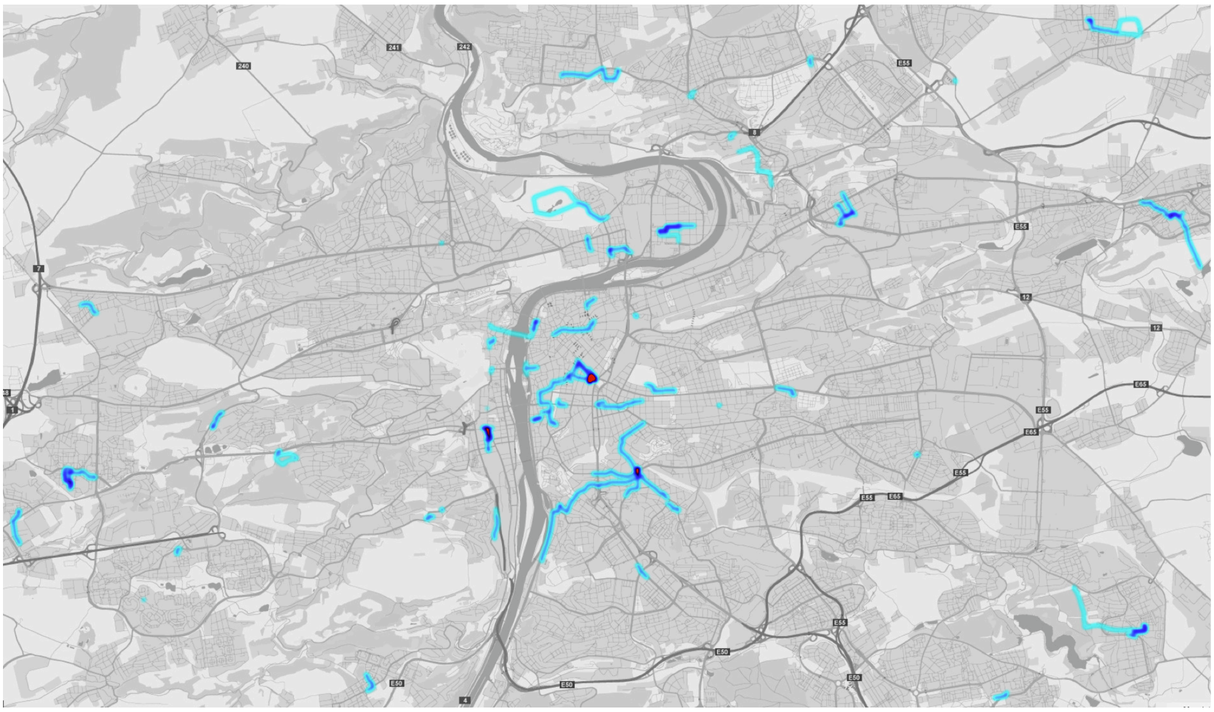


Figure 3.4: Heat map of routes in the whole city of Prague (ranging from blue – a low number of regular routes – to red – a high number of regular routes) (n = 20).

3.6 Implications

In this chapter, we showed how the verbal descriptions are shared among visually impaired people and that the amount of the information shared is remarkable. This fact suggests that there are an opportunity and meaning to support even unknown people in sharing their regular routes. Furthermore, in the results, we provide answers to RQ1 (How do people with visual impairments share spatial knowledge in a form of verbal route description?) and RQ2 (How much and how well do they remember their regular routes?) as defined in Chapter 1. Detailed answers to the research questions are summarized in the chapter Conclusion 8.

In the next chapter, we look at how exactly the sharing process proceeds when it is done in real time over the phone (tele-assistance). The motivation is to understand and gain in-depth insights into the strategies people with visual impairments take and to provide guidance for both professional tele-assistance centers but more importantly to individuals who wish to share their regular routes with other people with visual impairments.

Chapter 4

Blind-to-Blind Tele-Assistance

Abstract. *We raise a question whether it is possible to build a large-scale navigation system for blind pedestrians where a blind person navigates another blind person remotely by mobile phone. We have conducted an experiment, in which we observed blind people navigating each other in a city center in 19 sessions. We focused on problems in the navigator's attempts to lead the traveler to the destination. We observed 96 problems in total, classified them on the basis of the type of navigator or traveler activity and according to the location in which the problem occurred. Most of the problems occurred during the activities performed by the navigator. We extracted a set of guidelines based on analysis of navigation problems and successful navigation strategies. We have partially mapped the problem of tele-assistance navigation to POMDP based dialogue system.*

Note. *This chapter is based on published conference paper Navigation Problems in Blind-to-Blind Pedestrians Tele-Assistance Navigation [9].*

4.1 Introduction

Several approaches how to support independent traveling of visually impaired people exist. However, as mentioned in the previous chapter (Chapter 3) existing solutions are suffering from a lack of suitable data for route description. Further, in Chapter 3, we have shown that visually impaired people often navigate each other remotely by means of a phone call (tele-assistance) or emails and that this form of navigation is preferred over navigation by an untrained sighted person.

One already existing navigation solution is also based on a tele-assistance. It is a tele-assistance navigation center with professional navigators (see subsection 4.2.2). The main problem of this solution is its scalability as the gathering of the suitable set of landmarks for the particular area often requires the physical presence of the professional navigator on the spot. According to a study by Balata et al. [8], visually impaired people memorize relatively long routes at a very high level of detail. It was also shown that 67 % of visually impaired people have experience with sharing their route with friends/family via email, phone, messaging [10] and that they prefer navigation provided by a blind person to sighted person (also supported by [20], [60]). This opens the possibility to base a tele-assistance navigation service on visually impaired volunteers and to build up an efficient large-scale system where one visually impaired person navigates another. In this situation, the blind navigator (*navigator*) forms a natural source of suitable landmarks with their descriptions and with routing strategies optimized for blind travelers (*traveler*).

According to the functional model of a general navigation system for the blind defined by Loomis et al. [63], the *navigator* in such tele-assistance navigation service fully covers components providing a description of the environment (typically some kind of geographic information system), route planning, auditory display and speech input. The only component that cannot be covered independently by the *navigator* is the component responsible for determining the *traveler's* position and orientation. Here, collaboration with the *traveler* is needed. The *traveler* serves as a sensor gathering necessary data for the *navigator*, and can determine the position and orientation on her/his own.

The key feature of such tele-assistance navigation service is its non-stop availability. Here an automated dialogue-based navigation can be employed. There are several approaches to dialogue management: finite state machine, information state, grammar-based, plan-based and data-driven approach. Our case is highly complex, and thus data-driven approach like POMDP based dialogue managers is a suitable soulution [95].

Based on [62], [63], we identified the following five activities (three for *navigator* and two for *traveler*) that we wanted to observe in our experiment: The *navigator* describes the environment, plans the route (gives navigation instructions), and determines the blind *traveler's* position [63]. The *traveler* travels to a remote destination (executes navigation instructions), and senses the environment (identifies landmarks) [62].

Our main goal is to define a set of guidelines for efficient blind-to-blind tele-assistance navigation. We will investigate the process of tele-assistance-based navigation by blind people, with special reference to navigation problems that occur during this activity. Based on the analysis of the navigation problems we will develop recommendations for improving the training procedures in order to increase the efficiency of wayfinding in situations where tele-assistance takes place. Further, we will map the problem of tele-assistance-based navigation to POMDP based dialogue system in order to replace *navigator* with the computer system in the future.

4.2 Related Work

The related work section in this chapter follows the related work of previous chapters namely Chapter 1 and 3, and extends it with knowledge about pedestrian navigation.

4.2.1 Pedestrian Navigation

It has been shown that landmarks (representing landmark knowledge) are by far the most frequently-used category of navigation cues for pedestrians [67] (unlike junctions, distance, road type and street names or numbers). A study conducted by Ross et al. [88] states that the inclusion of landmarks within the pedestrian navigation instructions increased user confidence, and reduced or eliminated navigation errors. Rehrl et al. [85] showed that voice-only guidance in an unfamiliar environment is feasible and that participants preferred landmark-enhanced instructions.

The fact that humans rely primarily on landmarks to navigate from point A to B is reflected in many experimental designs of navigation systems, e.g., the system of Millonig and Schechtner [70]. The system designed by Hile et al. [53] presents a set of heuristics for selecting appropriate landmarks along the navigation path.

In our experiment, where the *navigator* instructs the *traveler* remotely without being physically present on the route and without any visual feedback, a declarative level of route knowledge is needed. The *navigators* were therefore thoroughly trained in compliance with official training methodology in the region where the tests were conducted [114]. The *navigators* were also introduced to objects that were not located on the test route. Finally, they checked a tactile map of the route and its environment to gain overview knowledge. In the training procedures for our experiment, we paid special attention to introducing all important landmarks and describing them to the *navigators* in order to support the creation of landmark knowledge (see Apparatus in section 4.3).

4.2.2 Orientation and Navigation of the Blind

In large spaces where body movement is necessary, visually impaired pedestrians use different cognitive strategies from those used by sighted pedestrians for navigation and orientation, based on egocentric frames [68], [69]. Typically, they have to memorize a large amount of information [102] in the form of sequential representation [69] based on routes. Route knowledge has to be acquired on a declarative level [47]. Fortunately, it seems that visually impaired people acquire superior serial memory skills. A study by Raz et al. [84] discovered that congenitally blind people are better than sighted people in both item memory and serial memory and that their serial memory skills are outstanding, especially for long sequences. In a study by Bradley and Dunlop [20], it was revealed that in a situation of pre-recorded verbal navigation, the blind navigator navigated the blind traveler significantly faster than a sighted navigator.

There are numerous navigation aids for visually impaired pedestrians. Some use special sensors to identify objects on the route, e.g., cameras [23], or an RFID based electronic cane [39]. Others are based on a concept described in [80] and rely on some kind of positioning system (e.g., GPS) in combination with the GIS system to identify objects and navigate the pedestrian, e.g., Ariadne GPS, BlindSquare. There have also been attempts to develop special interaction techniques for presenting navigation instructions, e.g., an auditory display [62] or a tactile compass [81].

The navigation aids based on major GIS systems (Google Maps, Apple Maps, OSM Maps, Nokia HERE Maps) suffer from an inappropriate description of the environment for visually impaired pedestrians. The available description may be imprecise (e.g., missing sidewalks or missing handrails), or may be ambiguous (e.g., an inadequate description of pedestrian crossing, meaning that it cannot be localized and identified without visual feedback) or it may ignore specific navigation cues (e.g., the surface structure of the sidewalk, acoustic landmarks such as the specific sound of a passage, the traffic noise of a busy street, or other sensory landmarks, such as the smell of a bakery). In addition, routing algorithms can encounter problems with non-trivial adjustments to the preferences and abilities of visually impaired people, e.g., their inability to cross open spaces (e.g., large squares).

Both inappropriate descriptions and unsuitable routing algorithms can be avoided by introducing navigation systems based on tele-assistance with a trained human agent. Various approaches have been proposed on the basis of various ways to identify the position and the environment of the pedestrian, like transmission of chest-mounted camera view to the navigator [23], a verbal description from the pedestrian optionally combined with GPS location and GIS [30], [100], or purely based on a verbal description and knowledge of the environment [109]. Namely Navigational Centre for the Blind [30], operating since 2007, proved to be helpful tele-assistance navigation service widely used (6650 cases in years 2008-2013 [29]) by the community of visually impaired people.

4.3 Experiment

In our experiment, we observed the process of navigation by a navigator navigating traveler by means of tele-assistance. The purpose of the experiment was to get insights into the process and to provide guidance for improvements (**RQ4**, Chapter 1). The navigation was performed in a way similar to that used in [20], [109]. The goal was to identify navigation problems in the following activities:

- *Navigator* describing the environment,
- *Navigator* giving navigation instructions,
- *Navigator* determining *traveler's* position,
- *Traveler* executing navigation instructions,
- *Traveler* identifying landmarks.

The experiment consisted of 19 sessions. There were two participants in each session, one in the traveler role and the other in the navigator role. Each session lasted around 100 minutes.

4.3.1 Participants

25 visually impaired participants (12 females, 13 males) were recruited via three methods: an e-mail leaflet sent to a group of Czech Blind United [30] clients, direct recruiting of our long term collaborators, and snowball technique. The participants in the experiment were aged from 25 years to 66 years ($mean = 43.44$, $SD = 13.27$). Fourteen participants had Category 4 vision impairment (light perception); 11 participants had Category 5 vision impairment (no light perception) [113]; 12 participants were congenitally blind, 13 participants were late blind. All of the participants were native Czech speakers. None of the participants in the *traveler* role knew the route before the experiment, though the character of environment was familiar to them. During recruitment, the participants were asked whether they are willing to participate in both roles, as the *traveler* in the first session, and then as the *navigator* in the following session. Table 4.1 contains details about the participants. Table 4.2 contains details about the sessions and about the roles that the participants took (the session IDs do not necessarily correspond to their real order). We tried to balance onset of impairment, category of impairment and gender of the participants in the sessions as much as possible. All of the participants (except P23) were

Part. No.	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**	**P14**	**P15**	**P16**	**P17**	**P18**	**P19**	**P20**	**P21**	**P22**	**P23**	**P24**	**P25**
Onset	C	L	C	C	C	L	C	C	C	C	L	L	L	C	C	L	L	L	L	L	C	L	C	L	L
Category	5	5	4	5	5	4	5	4	5	4	4	5	4	4	5	4	5	4	4	5	5	4	5	5	5
Gender	F	M	F	F	M	M	M	F	M	F	M	M	M	F	M	F	M	F	M	F	F	F	F	M	M
Age	38	60	37	42	29	61	43	27	32	29	36	37	50	38	43	38	60	50	62	66	25	65	29	31	58

Table 4.1: List of participants, including onset of the impairment (congenital – C, late – L), category of visual impairment [113], gender (male – M, female – F), and age.

Session	**S_1**	**S_2**	**S_3**	**S_4**	**S_5**	**S_6**	**S_7**	**S_8**	**S_9**	**S_{10}**	**S_{11}**	**S_{12}**	**S_{13}**	**S_{14}**	**S_{15}**	**S_{16}**	**S_{17}**	**S_{18}**	**S_{19}**
Traveler	P1	P3	P5	P6	P4	P8	P10	P11	P13	P14	P16	P17	P18	P19	P22	P21	P23	P24	P25
Navigator	P2	P4	P6	P3	P7	P9	P11	P12	P12	P15	P17	P18	P14	P20	P19	P22	P2	P23	P24
Duration	11:23	5:17	7:06	–	–	–	–	4:20	4:15	8:41	–	9:32	6:01	8:51	4:32	–	–	–	4:52
Success	Yes	Yes	Yes	No	No	No	–	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	No	No	Yes
Problems	1	1	4	8	11	7	4	0	1	3	11	4	4	7	3	4	3	12	8

Table 4.2: List of sessions, including participants' role in the experiment, duration of navigation (minutes), success of session, and the number of navigation problems.

active and regularly traveled alone. Several researchers have noted that it is quite difficult to acquire blind pedestrians as a target user group for a usability study [20], [23]. However, we had established a relationship with blind communities (namely Czech Blind United organization [30]) during our previous studies, and this made it comparatively easy to recruit a considerable number of blind participants for our experiment.

4.3.2 Apparatus

Training methodology

The goal of the training was to learn the *navigators* the route for regular independent walking, i.e., to form a declarative level of the route knowledge. We arranged several meetings with the chief methodologist from the Czech Blind United [30]. One of the chief methodologist's fields of expertise is in the training visually impaired people in spatial orientation and in preparing itineraries for their regular routes (i.e., routes to work, to a shop, to a public transport station/stop, etc.) in accordance with their navigation strategies. In order to conform with the official training methodology [114] used by the chief methodologist, we proceeded as follows: 1) We selected the route, identified important landmarks, and consulted possible dangers on the route together with the chief methodologist. 2) Together, we prepared a tactile map of the route and printed it on a paper using foil fuser technology (swell paper). 3) The experimenter observed the chief methodologist training the *navigator* in the first pilot session. 4) The experimenter trained the *navigator* according to the observed methodology under the supervision of the chief methodologist in the second pilot session. 5) The trained experimenter trained the *navigators* in all subsequent sessions of the experiment.

Description of the route

For our experiment, we selected a city center outdoor environment. Environments for this type of experiment are usually real environments [20], [85] rather than artificial (lab) environments,

Figure 4.1: Panorama from the beginning of the route containing decision points D1 and D2 (top), and from the end of the route containing decision points D3, D4, and D5 (bottom).

though exceptions are possible [93]. The location of the route was in a quiet area in the city center of Prague, Czech Republic (see Figures 4.1, 4.2). It was 256 m in total length (from S via D1-D5 to B11) and navigation via phone took place on the 105 m long final part of this route (from D1 to B11). In the initial part (from S to D1) of the route the *traveler* walked alone. This was done to allow the *traveler* to get oriented and to get familiar with the surrounding environment of D1. Along the route there were 5 decision points (D1-D5) to which the *navigator* tries to navigate the *traveler*, number of surface changes (SFx), acoustic landmarks (Ax), vertical traffic signs and columns (Cx), and doors (Bx) (see Figure 4.2).

Equipment

Our equipment and the data we collected is based on field laboratory design presented in Hoegh et al. [55]. The *traveler* was equipped with a Nokia 6120 mobile phone with a lanyard which hung from his/her neck. In this way, the phone was protected from being dropped unintentionally, and the *traveler* was able to release it and have an empty hand when needed, and s/he could also find it again quickly. The mobile phone was set to Czech language, and it was equipped with the MobileSpeak text-to-speech (TTS) screen reader application by CodeFactory. The *navigator* was located in the usability lab dedicated to executing user tests equipped with a laptop and the Skype application with Skype Out capabilities for connection with the mobile phone network. The laptop speakers and the internal microphone were used as input and output devices. Communication between the *traveler* and the *navigator* was recorded using MP3 Skype Recorder v3.1 (left/right channel separated for *traveler*/*navigator* communication).

In each session, we also recorded two video streams of the *traveler's* activities to be able to observe situation from different angles and to determine *traveler's* precise location. The first camera (GoPro Hero 3) recorded 1st person view and was installed on a shoulder strap of the backpack that was carried by the *traveler* during the session, while the second camera (Panasonic

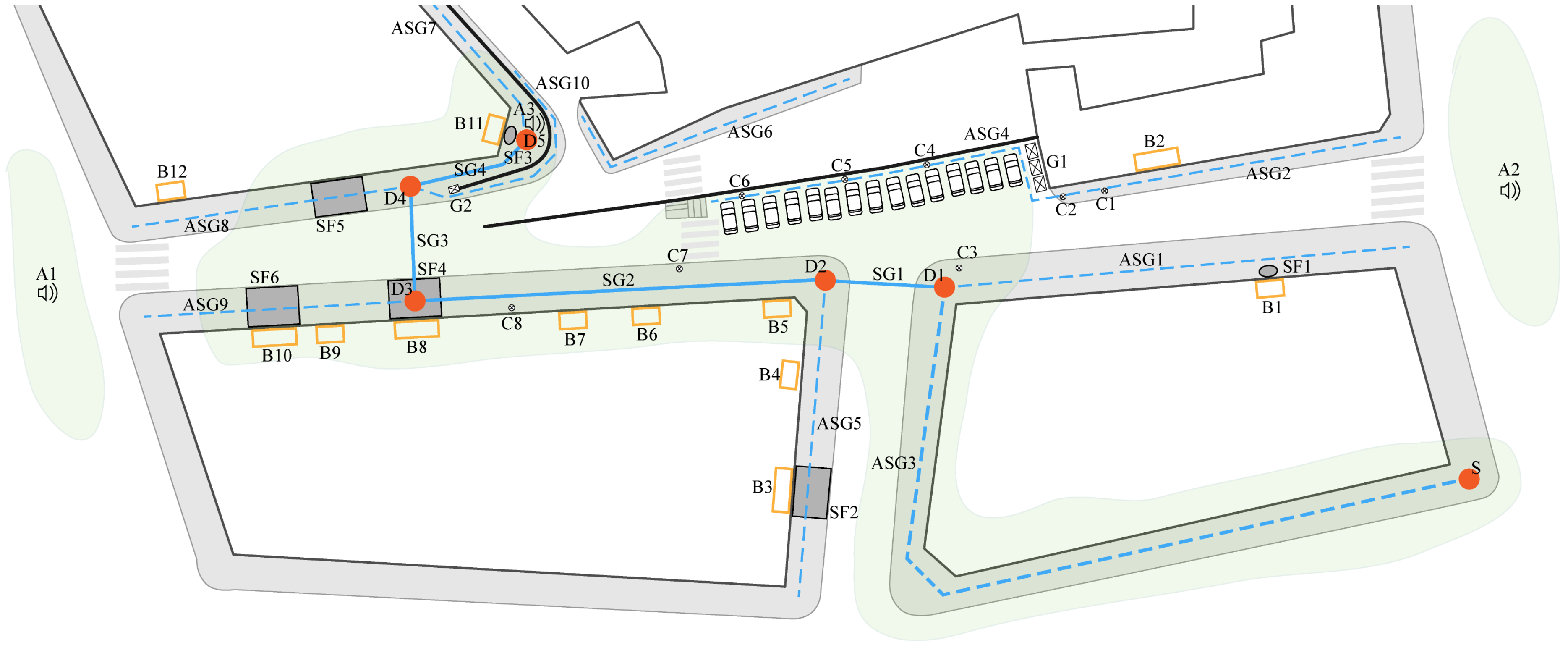


Figure 4.2: Schematic illustration of the route and its adjacent area. The rightmost red dot depicts the starting point of the route^{S}, red dots depict decision points^{Dx}, blue lines depict segments^{SGx} of the route (bold solid – route segments traveled with the *navigator*, bold dashed – adjacent segment^{ASG3} walked without the *navigator* from the starting point, dashed – adjacent segments ^{ASGx}), orange boxes depict buildings^{Bx}, dark grey boxes and ovals depict surface landmarks^{SFx} such as a rubber mat^{SF1,SF3} or differences in material, e.g., cobblestone^{SF2,SF4,SF6} and a broken sidewalk^{SF5}, speaker icons depict acoustic landmarks^{Ax}, such as busy streets^{A1,A2} or echoes^{A3}, crossed circles depict vertical traffic signs and columns^{Cx}, crossed squares depict waste containers of various sizes^{Gx}. The green area contains landmarks taught to the *navigator* in the training phase of the experiment. Bold black lines depict stone walls. Gray areas depict sidewalks.

SDR-S150) recorded a 3rd person view by the experimenter shadowing the *traveler*.

4.3.3 Procedure

Before the session started, both participants were briefed, and the purpose of the experiment was explained to them separately.

The experiment session consisted of two phases. In the first phase, the *navigator* was taught the route by the experimenter. In the second phase, the *navigator* navigated the *traveler* along the route. Both of the participants were asked to proceed as quickly and accurately as possible. The *traveler* was asked either to hold the phone in a hand or to leave it on the lanyard, according to their own preference.

The first phase of the experiment involved training the *navigator*. The training consisted of three walkthroughs. The first two walkthroughs of the route were done with the experimenter, and the third was done alone, with the experimenter in the vicinity. During the first two walkthroughs, the experimenter described the landmarks (see Figure 4.2) along the route and offered as many details as possible. During the third walkthrough, the *navigator* walked alone and asked the experimenter about the landmarks in cases when s/he was uncertain, so that s/he could remember better, but mostly to verify that s/he had learned the route sufficiently. After the walkthroughs, the *navigator* was accompanied into our usability lab and was presented with a tactile map of the overview of the route and the destination details. From this point, the *navigator* waited in the usability lab with the experimenter for a call from the *traveler*.

The second phase of the experiment consisted of a walkthrough of a part of the route by the *traveler*, and of navigation of the *traveler* by the *navigator*. This phase consisted of three parts. In the first part, the *traveler* was accompanied to the starting point of the route^{S} and was given the task. The task was given as follows: *"You have a meeting at Hostel Emma*^{B11} *(see Figure 4.2) on Na Zderaze Street. Now you are on the corner of Dittrichova Street and Resslova Street. Continue approximately 80 meters slightly downhill along Dittrichova Street to the first crossroad. The building will be on your right hand side, and on the left there will be cars parked on the sidewalk. Then turn right and continue approximately 80 meters uphill along Zahoranskeho Street to the crossroad with Na Zderaze Street. To reach the destination, you will have to call the navigator who knows the location very well. At the crossroad, you will be assisted with dialing the phone. Proceed as if you were alone, but we will be watching for your safety from a distance."* Then the *traveler* started out. The second part consisted of assisting the *traveler* with making the phone call from the corner^{D1} where the navigation with *navigator* starts. The phone call was initiated by the experimenter in the lab, who relayed the call to the *navigator*. Then the *traveler* accepted the phone call and started a dialog with the *navigator*.

The third part consisted of navigating the *traveler* by the *navigator* via a phone call. The *navigator* described the environment, gave navigation instructions, and determined the *traveler's* position. The *traveler* executed the navigation instructions, and identified the landmarks. The experimenter observed the whole session from nearby to ensure the safety of the *traveler*. If the *traveler* got lost beyond the possibility of finding the destination, and/or was in distress, the experimenter terminated the session. Otherwise, the *traveler* was not interrupted by the experimenter. After reaching the destination, the *traveler* was accompanied into the usability lab, where both participants were debriefed and received their payment.

4.3.4 Measures

During the sessions we measured the time to reach the destination in successful sessions and the number of navigation problems in all sessions. For the activities *Navigator* describing the environment and *Navigator* giving navigation instructions we define the navigation problem as deviation from the training *navigators* went through (see paragraph Apparatus – Training methodology). A navigation problem in the activity *Navigator* determining *traveler's* position is defined as concurrent occurrence of two events: the *traveler's* physical position differs from the *navigator's* imagination of the *traveler's* physical position, and the *navigator* is not determining the *traveler's* physical position. Problems in the *traveler's* activities (i.e., *Traveler* executing navigation instructions, *Traveler* identifying landmarks) are defined as fail to execute the navigation instruction and fail to identify the landmark.

4.3.5 Collected data

Nineteen Skype call audio files were recorded. Nineteen video files were recorded from a 3rd person view observing the *traveler*, and eighteen video files were recorded using a GoPro camera on the *traveler's* shoulder from the 1st person view (one file was not recorded, due to a hardware malfunction). These files were then merged and aligned by time into a single multimedia file for each session.

In order to analyze the data, we developed an application that allows time-stamped annotation of the *traveler's* physical position and the *navigator's* imagination of the *traveler's* physical position in the map. After annotation, the multimedia file from the session and both annotated positions in the map could be browsed side-by-side.

4.3.6 Results and Discussion

Eleven sessions finished with successful arrival at the destination after between 4 minutes 15 seconds and 11 minutes 23 seconds ($mean = 6\ min.\ 48\ sec.$, $SD = 2\ min.\ 28\ sec.$). One session was inconclusive due to intervention of the experimenter. Seven sessions failed. In this section, we describe the navigation problems observed during navigation of the *traveler* by the *navigator*. Initially, we focus on general results, and then we describe selected navigation problems in various types of situations.

We analyzed the navigation problems in all the sessions, and classified them into the corresponding activities performed by *navigators* and *travelers*, and into different types of situations. Out of the total of 96 problems, 71 were problems identified on the route, and 25 were problems identified off the route. Sixty problems were identified in the failed sessions, and 36 problems were identified in the successful sessions (see Table 4.2). Most of the navigation problems on the route (44 of 71) occurred in two activities: *Navigator* describing the environment, and *Navigator* giving navigation instructions (see Table 4.3). It was also shown that the *navigator's* problems on the route were clearly greater than the *traveler's* (53 vs. 18), but off the route there the difference was smaller (15 vs. 10).

A majority of the navigation problems that occurred in the activity *Navigator* describing the environment on the route were related to column^{C8} (8 out of 23) and decision point $[D3]$ (7

Situation (Landmark)	***Navigator* describing the environment**	***Navigator* giving navigation instructions**	***Navigator* determining *traveler's* position**	***Traveler* executing navigation instructions**	***Traveler* identifying landmarks**
			On the route		
Reorientation at a corner (D1, D2)	2	10	0	0	2
Crossing from corner to corner (SG1)	0	1	0	2	0
Traveling along a building (SG2, SG4)	4	4	2	3	1
Reorientation at a building (D3, D4)	0	3	1	0	0
Crossing from building to building (SG3)	0	1	0	1	0
Finding a landmark (B8, C8, SF3, SF4)	17	2	6	5	4
Total on the route	**23**	**21**	**9**	**11**	**7**
			Off the route		
Traveling along building (ASGx)	4	1	10	4	5
Finding landmark (B1)	0	0	0	1	0
Total off the route	**4**	**1**	**10**	**5**	**5**
All navigation problems	**27**	**22**	**19**	**16**	**12**

Table 4.3: Occurrence of navigation problems in the activities performed by *navigators* and *travelers* in different situations on and off the route.

of 23). *Navigators* did not mention column^{C8}, door^{B8}, and cobblestones^{SF4} while *traveler* was approaching decision point $[D3]$.

Along with the navigation problems in various types of situations (i.e., reorientation at a corner, crossing from corner to corner, traveling along a building, reorientation at a building, crossing from building to building, finding a landmark), we observed four other phenomena that affected successful navigation. They were: similarities in the environment; temporary changes of environment; landmark confusions; recovery from going astray.

Reorientation at a Corner

There were 14 navigation problems in reorientation at a corner. For example, the *navigator* did not instruct the *traveler* to turn left $[D1, S_{6,18}]$. The *navigator* did not relate the position of the *traveler* to the building $[D1, S_{16}]$. The *navigator* was unable to give the *traveler* unambiguous instructions on how to stand at the corner: *"Turn so you have the corner at your back."* $[D1, S_{13}]$. The *navigator* could not determine the *traveler's* position on the corner, i.e., which side s/he was on $[D2, S_{12}]$. The *navigator* wrongly instructed the *traveler* and confused *"turn left"* with *"have the building on your left"* $[D2, S_{11}]$. It seems that this situation was one of the most difficult for the participants.

However, several successful navigation strategies were used to reorient at the corner. For example, the *navigator* instructed the *traveler* to turn his/her back towards the building before s/he turned the corner $[D1, S_{15}]$. The *navigator* instructed the *traveler* to check if s/he could hear a busy street from the right $[D1, S_4]$. The *navigator* described the surrounding streets and gave their names at $[D1, S_8]$. The *navigator* checked on which side the *traveler* had a building and what the slope of the sidewalk was $[D2, S_6]$. The *navigator* asked the *traveler* on which side the downhill sidewalk was $[D2, S_{12}]$.

Crossing from Corner to Corner

We found 3 navigation problems when the street was crossed from corner to corner $[D1, D2]$. For example, the *traveler* did not execute the instruction to come back slightly from the corner to the street, so s/he arrived at the opposite corner, while the *navigator* expected him/her on the left from the opposite corner $[S_{11}]$. The *traveler* did not walk straight while crossing the street and missed the opposite corner $[S_{17}]$.

However, several successful navigation strategies were used to cross the street from corner to corner. For example, the *navigator* instructed the *traveler* to return back to the street and cross, in order not to miss the corner on the other side $[S_{10,17}]$. The *navigator* instructed the *traveler* to walk around the cars from the left side in order not to miss the corner $[S_{12}]$. The *navigator* instructed the *traveler* to cross the street to the opposite sidewalk. In this way, the *navigator* used previously traveled route of the *traveler* and the fact, that the street had sidewalk on both sides, for giving the instruction $[S_8]$.

Traveling along a Building

In the situation when traveling along a building, we observed 14 navigation problems. For example, the *traveler* did not describe the slope of the sidewalk precisely when the *navigator* was trying to determine his/her position $[SG2, S_{16}]$. The *traveler* did not execute the instruction to walk along the building $[SG2, S_7]$. The *navigator* did not know about two restaurants^{B6,B7} that the *traveler* asked about $[SG2, S_7]$. The *navigator* did not instruct the *traveler* to walk along the building in order to find the rubber mat^{SF3} $[SG4, S_{3,4,5,15,19}]$.

However, several successful navigation strategies were also used for traveling along building. For example, the *navigator* described the sidewalk made of small paving blocks $[SG2, S_{3,14}]$. The *navigator* checked that the building was on the left-hand side of the *traveler* $[SG2, S_{7,8}]$. The *navigator* checked the sound from the busy street in front of the *traveler* $[SG2, S_{14}]$. The *navigator* described the restaurants on the left-hand side $[SG2, S_{3,8}]$.

Reorientation at a Building

We found 4 navigation problems during reorientation at a building. For example, the *navigator* did not determine the position of the *traveler* when s/he reached the doors^{B8} and instructed him/her to turn right instead of instructing him/her to turn about face when s/he was facing the door^{B8} $[D3, S_3]$. The *navigator* did not determine the *traveler's* orientation when s/he reached the other side of the street $[D4, S_3]$.

However, several successful navigation strategies were also used for reorientation at the building. For example the *navigator* checked that the building was on the left-hand side of the *traveler* after s/he had crossed the street $[D4, S_{3,8,9}]$. The *navigator* instructed the *traveler* to have the doors^{B8} behind his/her back $[D3, S_4]$.

Crossing from Building to Building

Two navigation problems were observed during crossing the street from one building to another building $[D3, D4]$. For example, the *navigator* instructed the *traveler* to turn right if s/he found cars parked along the sidewalk, instead of bypassing them [S_3]. The *traveler* did not execute the instruction to cross the street to the building, and stopped at the edge of the sidewalk [S_{15}].

However, several successful navigation strategies were also used for crossing from building to building. For example the *navigator* instructed the *traveler* to walk around the parked cars from the left [S_1].

Finding a Landmark

In the situation of finding a landmark, we observed 34 navigation problems. For example, the *navigator* did not describe the column^{C8} [$S_{2,4,12,13,14}$]. Similarly, the *traveler* did not identify the same column^{C8} even if s/he struck it [S_{14}]. The *navigator* did not describe the wooden doors^{B8} with metal fittings and a handle at head level [$S_{10,11,14}$]. The *traveler* failed to check the material of the doors^{B8} and the handle [$S_{15,18}$]. The *traveler* did not execute the instruction to stop at the cobblestones^{SF4} although s/he did find them [S_{19}]. Alternatively, the *traveler* did not identify the cobblestones^{SF4} at all [$S_{10,12}$].

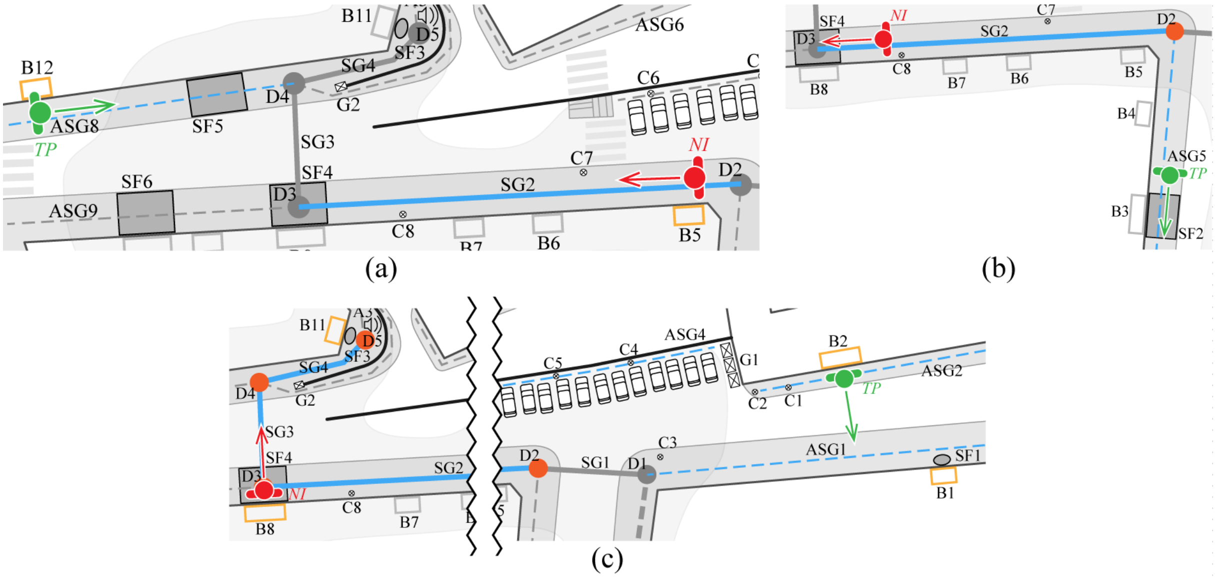
However, several successful navigation strategies were also used for finding a landmark. For example, the *navigator* described the cobblestones^{SF4} [$S_{5,13}$]. The *navigator* described the wooden door^{B8} with metal fittings and a handle at head level [$S_{1,4,12}$]. The *navigator* described exact position of the column^{C8} – 15 cm from the building on the left side [$S_{5,7,8,9,18}$]. The *navigator* described the distance to the column^{C8} from the corner [$D2, S_{8,9,19}$]. The *navigator* described the acoustics^{A3} at the corner [$D5, S_{12,13,18,19}$]. The *navigator* described a rubber mat^{SF3} on the sidewalk [$SG4, S_{12,14,18}$].

Similarities in the environment

If the *traveler* was inattentive to the details of landmarks, two parts of the route can seem to be very similar. The similar parts can be characterized by the same sequences of similar landmarks (e.g., route part R consists of landmarks A, B, C and route part R' consists of landmarks A', B', C', where A is similar to A', B to B', and C to C').

There was similarity between one sidewalk^{ASG8} from restaurant^{B12} to place with broken sidewalk^{SF5} and another sidewalk^{SG2} from shop^{B5} to the cobblestones^{SF4} [S_5] (see Figure 4.3(a)). The *navigator* thought that the *traveler* had crossed the street and returned back (from SG4 back to the other side of the road to SG2 and farther away to D2), as they could not find the destination^{B11}. This was because of incorrect instructions from the *navigator* – s/he did not stress that the *traveler* should go along the building to find the rubber mat^{SF3} at the destination [$SG4$]. The *navigator* checked the acoustic landmark^{A2} and the *traveler* acknowledged that there was indeed a busy road^{A1} behind his/her back; however, it was the other one^{A2}. They did not check the material of sidewalk: on one sidewalk^{SG2} there are small paving blocks, whereas the other one^{ASG8} there is asphalt.

There was similarity between one sidewalk^{ASG5} from corner^{D2} to cobblestones^{SF2} and another sidewalk^{SG2} from corner^{D2} to cobblestones^{SF4} [S_{11}] (see Figure 4.3(b)). The *navigator* confused the navigation instruction (left vs. right), and the *traveler* continued to the left^{ASG5} instead of to the right^{SG2} [$D2$]. The *navigator* did not check whether the *traveler* had buildings on his/her left side, and did not check which side the landmarks reported by the *traveler* were



(a) (b)

(c)

Figure 4.3: Similarities in sessions S_5 (a), S_{11} (b), and S_{18} (c). The green figure represents the *traveler's* physical position and direction (TP). The red figure represents the *navigator's* imagination of the *traveler's* physical position and direction (NI).

on. Both sidewalks^{ASG5,SG2} are downhill, but the first^{ASG5} is much steeper. The *navigator* checked for the slope and the *traveler* acknowledged that it was downhill but not how steep it was. Navigation continued until the *traveler* reached cobblestones^{SF2}. The material of the doors^{B3} did not match the right one^{B8}, however the *navigator* instructed the *traveler* to continue further downhill^{ASG5}.

There was similarity between one sequence of sidewalks^{ASG4,ASG2,ASG1} from column^{C4} near stone wall^{ASG4} to doors^{B1} with rubber mat^{SF1} and another sequence of sidewalks^{SG2,SG3,SG4} from corner^{D2} to doors^{B11} with rubber mat^{SF3} [S_{18}] (see Figure 4.3(c)). The *navigator* forgot to turn the *traveler* to the left to cross the street^{SG1} and the *traveler* ended up by stone wall^{ASG4}. The *traveler* continued along the wall^{ASG4} and reported cars along the buildings. The *navigator* acknowledged, but did not check how far from the building the cars were. There is no sidewalk, there are no buildings, but there is a stone wall. The *traveler* reported wooden doors^{B2} but the *navigator* did not check for cobblestones^{SF4}, which are missing there^{B2}. After crossing the street from the doors^{B2}, the *traveler* did not find the corner on the right side and decided to walk to the left in the opposite direction^{ASG1}, but the *navigator* did not make any comment. The *traveler* reported that s/he was at the destination on the at doors^{B1} with rubber^{SF1}.

Landmark confusion

Travelers often had to make a further examination of a landmark that they had discovered, in order not to confuse it with another object. The *traveler* confused a railing with a temporary traffic sign placed next to a column^{C3} [S_1]. The *traveler* confused a garbage container^{G1} with a trash can^{G2} [S_5]. The *traveler* confused cars with garbage containers^{G1} [S_{18}]. In session The *traveler* confused a passage (there was change in acoustics – an echo) with a van parked along the sidewalk [$D3, S_{19}$]. *Travelers* confused a building with a stone wall^{ASG4} [$S_{6,16,18}$].

Temporary Changes of Environment

An urban environment has a rhythm of its own. Streets that are busy during the day are silent at night, and the shops are closed. Some shops open at 9:00 am, while some restaurants and coffee bars do not open until 11 o'clock. There is also a weekly rhythm of dustmen and periodic street cleaning. All of these changes have an impact on the environment and affect some of the landmarks. In our case, there was increased traffic in an otherwise quiet street near $[D1, S_{14}]$, a temporary traffic sign next to column^{C3} $[S_1]$, a dustbin put outside a door^{B8} for the garbage collectors $[S_7]$, a van parked on the sidewalk $[D3, S_{19}]$, or a missing advertising stand $[D2, S_{16}]$.

However, none of these temporary changes of environment caused any serious navigation problem in our experiment except for session S_{16} (the *navigator* wanted the *traveler* to find the advertising stand and sent him/her in wrong direction).

Recovery from Going Astray

During the sessions, we observed *traveler* going astray. In a moment when the *navigator* or the *traveler* realized that the *traveler* is out of the route they started to recover from this situation and get back on the right route. The recovery process can be divided into three subsequent steps: 1) realize that the *traveler* went astray, 2) determine the *traveler's* real position, 3) take the *traveler* back on the route.

The *traveler* walked relatively long time without mentioning that s/he went astray. This especially happened when the *navigator* did not determine the *traveler's* physical position on regular basis and when the *traveler* failed to identify the landmarks. For example, at second cobblestones^{SF6}, the *traveler* realized that s/he probably did not identify the first ones^{SF4} as the door^{B10} material did not match the *navigator's* description $[S_4]$. However, neither the *traveler* nor the *navigator* realized that the *traveler* went astray, but they were convinced that they reached the destination^{B11}, even though they were at different doors^{B1} $[ASG1, S_{18}]$.

In order to take the *traveler* back on the route the *navigator* had to determine the *traveler's* position and direction. For example, when the *navigator* realized the *traveler* went astray, s/he asked: *"Where can you hear the loud cars? Theoretically, from the front left?"* when the *traveler* confirmed the *navigator* determined his/her position at stone wall^{ASG4} near 3rd column^{C4} $[S_6]$. However, neither the *navigator* did not determine the *traveler's* position until the experimenter terminated the session $[S_{4,5}]$.

The last step is the attempt to take the *traveler* back on the route. This was typically done by backtracking to last known point on the route. For example, the *navigator* successfully instructed the *traveler* to cross the street from stone wall^{ASG4} near 3rd column^{C4}, *"Try to cross the street straight. You need to get to the curb."* $[S_6]$. The *navigator* successfully instructed the *traveler* to return back from end of the street^{ASG9} back to first cobblestones^{SF4}, *"You have to go back and have houses on the right."* $[S_{12}]$. However, the *traveler* was not able to cross the street from stone wall^{ASG4} and the session was terminated on his/her request $[S_{17}]$.

After successful recovery they tried to identify the error either in giving the navigation instruction by the *navigator* or in executing the navigation instruction by the *traveler*. The most common error was not identifying a cobblestones^{SF4} by the *traveler*. The common solution was returning back and trying to properly identify the landmark $[S_{4,12,14}]$.

Guidelines

We have defined following five guidelines for efficient blind-to-blind tele-assistance navigation based on the analysis of 96 navigation problems and successful navigation strategies collected during the experiment (see Table 4.4).

	Activity	**Guideline**	**Example (situation)**
G1	*Navigator* describing the environment	*Navigator* should describe the environment in detail, with focus on tactile properties (materials, changes of materials, slopes) and auditory properties (traffic sounds, echoes).	✓ The *navigator* described the sidewalk made of small paving blocks (4.3.6). ✓ The *navigator* described the surrounding streets and gave their names (4.3.6). ✗ The *navigator* did not describe the wooden doors (4.3.6).
G2	*Navigator* giving navigation instructions	*Navigator* should relate the orientation of *traveler* to the environment, to *traveler's* previous route, and/or to auditory landmarks; *navigator* should describe landmarks along the route, mention on which hand side is the leading line (building, edge of sidewalk), and mention auditory properties.	✓ The *navigator* instructed the *traveler* to have the doors behind his/her back (4.3.6). ✓ The *navigator* instructed the *traveler* to walk around the parked cars from the left (4.3.6). ✗ The *navigator* did not relate the position of the *traveler* to the building (4.3.6).
G3	*Navigator* determining *traveler's* position	*Navigator* should regularly check *traveler's* position, e.g., ask about execution of instruction and discovered landmarks.	✓ The *navigator* checked that the building was on the left-hand side of the *traveler* (4.3.6). ✓ The *navigator* checked the sound from the busy street in front of the *traveler* (4.3.6). ✓ The *navigator* checked on which side the *traveler* had a building and what the slope of the sidewalk was (4.3.6).
G4	*Traveler* executing navigation instructions	*Traveler* should listen to the whole navigation instruction before execution, restate instruction, acknowledge both understanding and execution of the instruction.	✗ The *traveler* did not execute the instruction to come back slightly from the corner to the street (4.3.6). ✗ The *traveler* did not execute the instruction to cross the street to the building, and stopped at the edge of the sidewalk (4.3.6). ✗ The *traveler* did not execute the instruction to walk along the building (4.3.6).
G5	*Traveler* identifying landmarks	*Traveler* should describe the environment in detail, with focus on tactile and auditory properties.	✗ The *traveler* failed to check the material of the doors (4.3.6). ✗ The *traveler* did not describe the slope of the sidewalk precisely when the *navigator* was trying to determine his/her position (4.3.6). ✗ The *traveler* did not execute the instruction to stop at the cobblestones although s/he did find them (4.3.6).

Table 4.4: Guidelines for efficient blind-to-blind tele-assistance navigation extracted from observed navigation problems and successful navigation strategies, their application, and some of the examples from the experiment.

POMDP Based Dialogue System

The findings obtained in the study can be used for POMDP based dialogue system definition [22], [120]. In our case the system is represented by the *navigator* and environment is represented by the *traveler*.

A POMDP is defined by sextuplet $<S, A, Z, T, O, R>$, where S is a set of states (*Traveler's* states), A is a set of the system's actions (*Navigator's* actions), Z is a set of observation the system can experience (a set of *Navigator's* observations), T is a transition model, O is an observation model, and and R is a reward model.

The state set S=$<I_t \times P_t \times D_t>$ is composed of three features: *traveler's* action (I_t), which corresponds to *Traveler* executing navigation instruction activity, *traveler's* 2D coordinate position (P_t), and *traveler's* direction (D_t) as an absolute angle. In the future the state set can be extended with features such as type of *traveler's* disability or his/her experience. The action set A=$<I \times L>$ is composed of two features: action (I), and landmarks (L). The actions are passed to *traveler* during *Navigator* giving navigation instruction activity. The observation set Z=$<OI_t \times OL_t \times OD_t>$ is composed of three features: observed *traveler's* action (OI_t), observed *traveler's* landmarks (OL_t), which corresponds to *Traveler* identifying landmarks activity, and observed *traveler's* direction (OD_t) as a relative angle. The observation set is acquired during *Navigator* determining *Traveler's* position activity. To parametrize the transition model $T(s', s, a) = p(s'|s, a)$ we can use *Traveler* executing navigation instruction activity. In the future the transition model can be used for personalisation based on types of disability or experience (from the state set) such as adjustment of segment length, or usage of specific landmarks. The observation model $O(s', a, z) = p(z|s', a)$ is represented by GIS-like data structure with probable *traveler's* states. Findings from sections 4.3.6, 4.3.6, and 4.3.6 can be used for parameterization. In the future the observation model can be used for probability distribution visualization of *traveler's* position. In the future the reward model $R(s, a) : S \times A \to \mathbb{R}$ can be parameterized by stress function e.i. whether to re-plan or back-track when the *traveler* went astray and his/her stress level became high (see section 4.3.6).

4.4 Conclusion

We have gathered a set of problems that occur during the process of blind-to-blind navigation by means of tele-assistance. These problems have been classified into activities performed by the *navigator* and by the *traveler* and have been assigned to categories of situations where these problems occurred. We have also described in detail behavior of *navigators* and *travelers* in special situations (i.e., similar parts of the route, temporary changes in the environment, landmark confusion, and recovery from going astray). It seems that substantial number of problems are related to activity *Navigator* giving navigation instructions. Based on these finding we have defined five guidelines for blind-to-blind tele-assistance navigation. Further, these findings and guidelines can serve as a basis to improve the training for visually impaired people to make the wayfinding process more efficient in situations when tele-assistance takes place. Furthermore, our results are suitable for parametrization of POMDP based dialogue systems, which can form a step towards replacement of human *navigator* by a computer system.

Future research should focus on experiments in different environments (e.g., city park, indoors) and on the development of efficient training methods for blind-to-blind pedestrian tele-assistance based navigation.

4.5 Implications

In this chapter, we provided a set of guidelines for tele-assistance navigation both for sighted and visually impaired navigators. We categorized the successful strategies and drafted a model of the dialog system, which could autonomously navigate pedestrians with visual impairments. However, the amount of data collected which is insufficient at this time for training of POMDP based system. Furthermore, we provide answers to RQ3 (How do people with visual impairments perform while sharing the verbal route description over the phone?) and RQ4 (How can be tele-assistance navigation for people with visual impairments improved in general?) as defined in Chapter 1. Detailed answers to the research questions are summarized in the chapter Conclusion 8.

In the next chapters, we utilize the findings from this chapter in two ways. First, we describe a theoretical collaborative navigation system enhanced by game elements (Chapter 5) based on our tele-assistance guidelines. Second, we utilize transcribed verbal description, strategies and course of the dialogs to implement and model a conversational agent for navigation (Chapter 6 focuses on natural language generation, Chapter 7 describes the model of conversational agent and prototype of the implementation).

Chapter 5

Collaborative Tele-assistance Navigation

Abstract. *This chapter shows how elements of gamification, i.e., game thinking and game mechanics, can be integrated into a collaborative navigation system for visually impaired persons in order to encourage them to travel independently and thus to improve their quality of life and self-confidence. The navigation system supports independent navigation in unknown places by mediating help from another visually impaired person, who is familiar with the particular place. The system has been successfully enhanced by game elements and illustrates the potential of introducing game elements into these kinds of systems.*

Note. *This chapter is based on published book chapter Game Aspects in Collaborative Navigation of Blind Travelers [15].*

5.1 Introduction

In the last chapter, we showed how people with visual impairments could navigate each other over the phone, and we presented a set of guidelines, which should make the process more efficient. Further, in the Chapter 3, we showed that the length of regular routes people with visual impairments walk is substantial and that they to some extent share their verbal descriptions with friends and family.

These findings led us to base a tele-assistance service on visually impaired volunteers and to build up a concept of a system in which one visually impaired person navigates another one. To reach a satisfactory level of efficiency, we need to solve one problem – we need to reach a critical mass of volunteers who will cover a significant number of routes and to ensure almost non-stop availability of the service. We focus on the problem of recruiting volunteers, which can be approached by introducing game elements into the navigation system.

In this chapter, we illustrate how game elements can be integrated into a navigation system based on a collaborative tele-assistance service. The introduction of game aspects is aimed at attracting a critical mass of volunteers, and at supporting the social activity of visually impaired people through cooperation within the framework of a gamefied navigations.

5.2 Related Work

The related work section summarizes research on serious games, for further reading about people with visual impairments see Chapter 1.

5.2.1 Gamification and Games

Gamification can be defined as “the use of game design elements in non-game contexts,” [34]. According to the Oxford Dictionary, gamification is defined as the application of typical elements of game playing (e.g., point scoring, competition with others, rules of play) to other areas of activity. One of the main ideas of gamification is to make applications attractive to users, and in this way to increase their motivation to solve some specific problems. This feature is particularly important in the field of healthcare. It is necessary to motivate the patient to cooperate during a specific treatment. This motivation can be reinforced by including game elements into a particular application that is used during treatment. In our case, a suitable navigation system can help during the sixth phase - coping and mobilization - of the adjustment process described in the Introduction. Gamification of this kind of navigation system can increase its therapeutic effect. The navigation system will be more intensively used, and other competencies can be improved, e.g., social communication in the competitive or cooperative contests introduced by the game. This idea is supported by the conclusions drawn by Sawyer [92], who states that games can help in therapeutic activities from a personal perspective.

Game development requires the fulfillment of essential factors of the game. In the book Serious Games, by C. C. Abt [1], we can find the following definition of a game, which in part fits to our context: “A game is an activity among two or more independent decision-makers seeking to achieve their objectives in some limiting context.” However, this definition leaves out the cooperative aspect of a game, making it too narrow for our purposes. Another definition, given by Salen and Zimmerman [90], is as follows: “A game is a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome.” The system is represented by four elements: objects, attributes, internal relationships, and environment. The system can be closed or open (there is an exchange of some kind with its environment). A game can be framed as a formal, experiential, or cultural system. Players interact with the system in order to experience playing the game. The conflict in games is artificial and can be described as a contest between players, which can take many forms, from cooperation to competition. There are rules that limit player behavior and define the game. Each game has a quantifiable goal or outcome. We will describe our collaborative navigation system in terms of this definition in a later section.

5.3 Collaborative Navigation of Visually Impaired

The navigation and orientation strategies of visually impaired people described in the literature make one fundamental assumption: that the person knows the route to the desired destination. The route can be described as a set of step-by-step instructions and a description of the environment where the navigation takes place. This specific environment description, along with step-by-step instructions, can be obtained in several ways and at various levels of detail.

The least detailed description can be obtained from existing map sources. This is often limited to a sequence of direction information and approximate distances. A sufficient description of the environment can be obtained using a navigation instructor (either via phone or in person). There are tele-assistance centers [30], from which a navigation instructor can describe a part of the route to a visually impaired person. The descriptions are usually prepared in advance and are stored as text files. In many cases, the navigation instructor does not have direct experience with the particular environment.

The best option for getting an environment description is to arrange a meeting with a navigation instructor in person. Then the navigation instructor teaches the visually impaired person his/her desired route (e.g., the route to a shop, to work, to the doctor) [114]. The process of learning a new route (obtaining an environment description) usually consists of the following steps: first, the navigation instructor leads the visually impaired person through the route and describes his/her surroundings (this part is done twice). Next, the visually impaired person walks alone and is supervised by the navigation instructor (this part is also done twice). Finally, the visually impaired person can get the description in printed form, optionally with a haptic map. This method is regularly used and the description that is learned is detailed and usable, but it costs a lot of time and money.

It is clearly essential to have a detailed description of the environment where the navigation takes place. However, this description cannot be obtained automatically and has to be created manually by or with the help of a navigation instructor. This is a common problem for all navigation systems for the visually impaired. There is also clearly a problem with the small number of instructors and with their availability. Our navigation system prototype, therefore, aims to find ways to utilize the environment descriptions and regularly walked routes already known by visually impaired people.

As already described in Chapter 3, a description of an environment can also be acquired by visually impaired persons themselves, though with limitations due to their level of impairment. As a reminder, we list typical landmarks acquired by visually impaired persons themselves are:

- leading lines formed by edges of the pavement, handrails, corners of buildings, the better side of the street to travel on, etc.,
- sounds from traffic, construction work, water, the width of the street from the echo, etc.,
- smells of various types, e.g., bakeries, drugstores, sewers, etc.,
- the direction of heat from the sun at certain times of day,
- approximate distances, changes in elevation, type of ground material.

The names of streets, the order of turns, directions through parks and large open areas, or the location of a certain house in the street have to be learned with all the useful landmarks by a navigation instructor (orientation and mobility specialist).

In the Chapter 3, we showed that visually impaired people remember a detailed description of regularly walked routes and they create their own mental models of the environment, consisting of a large number of landmarks. They are able to help other visually impaired people to create their own mental models, e.g., by describing navigation instructions step-by-step through a known route. Although there are limited numbers of visually impaired people able to provide descriptions of their regular walks, our previous research has shown that the length of regularly walked routes is approximately 5 km per person [8]. This length seems to be sufficient for our purposes. What is necessary is to form a pool of people who, in combination, know many routes in a city. In this way, it is possible to provide adequate route descriptions for a relatively large community of visually impaired persons.

The idea of sharing descriptions of regular routes with other visually impaired people arises from current practice. Visually impaired people share some of their routes with friends and family via e-mail or instant messaging. A typical use-case is visiting a friend: a host sends an e-mail to a visitor with a detailed description of the route from a public transport stop to his/her home. Balata et al. [8] also investigated the willingness of the visually impaired people participating in the study to share their routes with other unknown people in order to help them with navigation. They found that 80 % of participants were willing to share their routes.

The lengths of regularly walked routes and the current custom of sharing known routes with friends and family enable us to define two essential conditions for a navigation system based on collaboration among visually impaired people (for more details see Chapter 3):

- sufficient length of regularly walked routes,
- the ability and willingness to share regularly walked routes with other people.

Unlike state-of-the-art navigation systems for visually impaired persons based on a description of the environment, our system draws on knowledge already distributed among the user group of visually impaired people. In this way, we can address the issue of insufficient numbers of navigation instructors and the fact that sufficiently detailed environment descriptions cannot be generated automatically.

5.3.1 Regular Route

In this subsection, we define the term regular route. A regular route is typically the route to work, to the local shopping area, to the nearest public transport stop, or a route that a visually impaired person uses for taking a walk at least once a week. S/he knows every detail of the regular route, e.g., the optimal leading lines, cracks in the pavement, smell and sounds (echoes), strategies for crossing a street, etc. Our study [8] shows that these routes are approximately 5 km in length per visually impaired person (for more information see Chapter 3).

These routes can be collected by tracking the user of the collaborative navigation system or can be collected via an interview and plotted into a geographical database. These two methods can be used to produce a map showing the density of regular routes, which can be rendered as a heat map (see Figure 5.1). Areas with a greater density of regular routes have a greater number of visually impaired people regularly walking through that area so that higher-quality environment descriptions are available for navigation.

5.3.2 Use-case

Let us imagine a visually impaired woman, Alice, who is on her way to a music club located near Bob's home on a Saturday night. Bob is also visually impaired, but he does not know Alice.

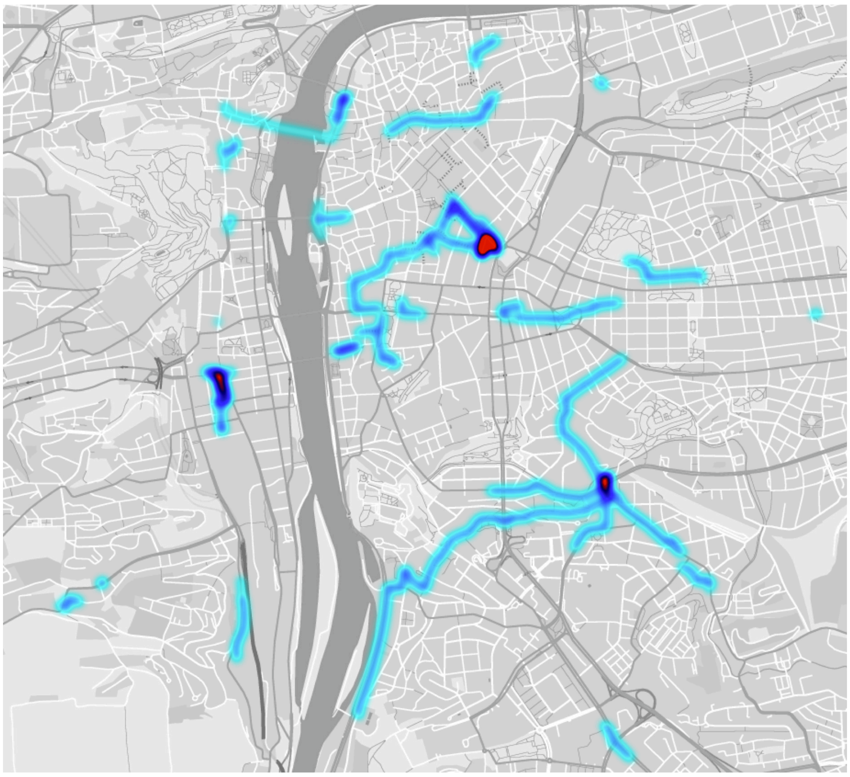


Figure 5.1: Heat map showing regular routes (ranging from blue – low number of regular routes – to red – high number of regular routes available) (for more details about the data collection see Chapter 3) (n = 20).

Although she has prepared a route itinerary generated by Google Maps walking directions, she is nervous because she has never walked the route before. At a complicated crossroads, she does not know how to cross the street. She decides to use collaborative navigation system and is connected via a phone call to Bob, who passes the crossroads every day. She describes her traveled route and her destination, and she describes her surroundings. She agrees with Bob on her precise location, and he gives her a detailed description of the route to the music club. Finally, she reaches the club, as planned (see Figure 5.2).

5.3.3 User Roles

From the use-case above we can define two basic user roles of the collaborative navigation system for visually impaired users – *traveler* and *navigator*. However, there is one more user role absent from the use-case, which is as important as the first two – *explorer*. The following paragraphs describe user roles and transitions between them in detail.

Traveler

A *traveler* is a visually impaired person traveling independently along an unknown route (in our use-case, it is Alice). If s/he is uncertain about the next steps in his/her navigation s/he can request help in navigation help from the Navigator. The *traveler* uses the system to connect to the *navigator* and then, via dialog, they solve the navigation problem (the Navigator listens to the *traveler's* route and surroundings, and they find and agree on an unambiguous orientation

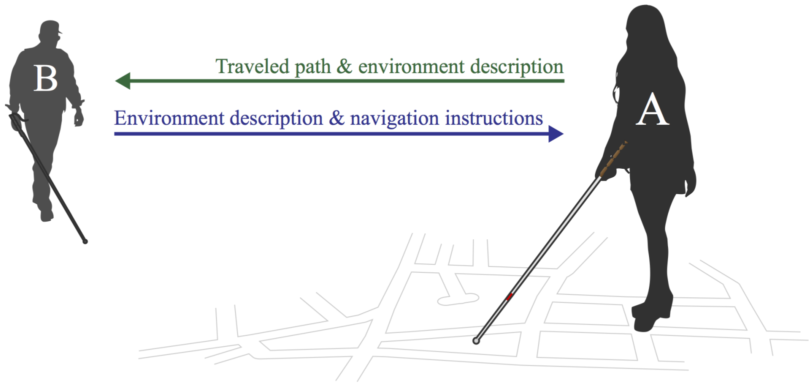


Figure 5.2: Person A (Alice) describes her traveled route and the surrounding area. On the basis of her description, person B (Bob) localizes person A and gives her navigation instructions.

point. The *navigator* gives navigation instructions to the *traveler* [109], [9].

Navigator

A *navigator* is a visually impaired person (in our use-case it is Bob) who knows several regular routes. The *navigator* is available to give navigation advice on his/her regular routes, and s/he can remotely (from another part of the city, or from his/her home) navigate the *traveler* (in our use-case it is Alice) via a phone call. The *navigator* selects days and times when he will make himself available to be a *navigator*.

Explorer

An *explorer* is a visually impaired person with an adventurous spirit. The *explorer* likes to explore new routes and new venues in order to be able to navigate *travelers*. After a period of time, when the *explorer* has learned a new route, s/he becomes a *navigator* on the new route.

If the *explorer* gets lost while exploring, s/he can change his/her user role and become a *traveler*. In this role, s/he can use the system to get navigation aid from a *navigator*, if there is one available. If the *explorer* learns a new route sufficiently (if s/he is able to navigate another visually impaired user of the system on this route), s/he changes his/her user role and becomes *navigator* for this route. If a *navigator* gets lost on an unknown route, s/he changes his/her user role and also becomes *traveler*. In this role, s/he can then use the system to get navigation aid from another *navigator* for this route (see Figure 5.3).

There is typically more than one *navigator* for a given route (see Figure 5.4). When this is the case, a *navigator* is selected on the basis of his/her availability, success rate (number/proportion of successful navigation attempts), occupancy (number of navigation attempts within a certain period of time), rating (subjective evaluations submitted by navigated *travelers*), and several other parameters, e.g., level of impairment, etc.

5.3.4 Implementation

There are several ways to implement the collaborative navigation system. There are in fact two problems to solve. The first is to collect the data that will form a regular database of routes.

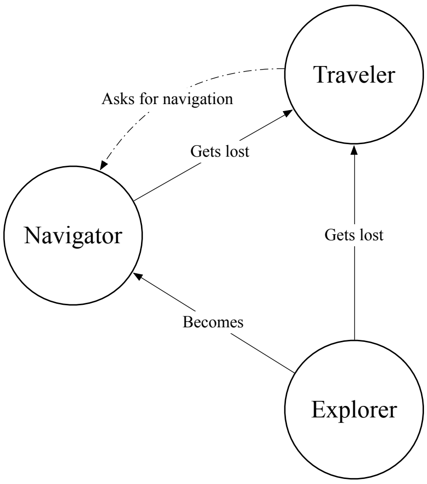


Figure 5.3: A diagram of the collaborative navigation system with transitions in user roles (depicted by a solid line) and use-cases (depicted by a dash-dotted line).

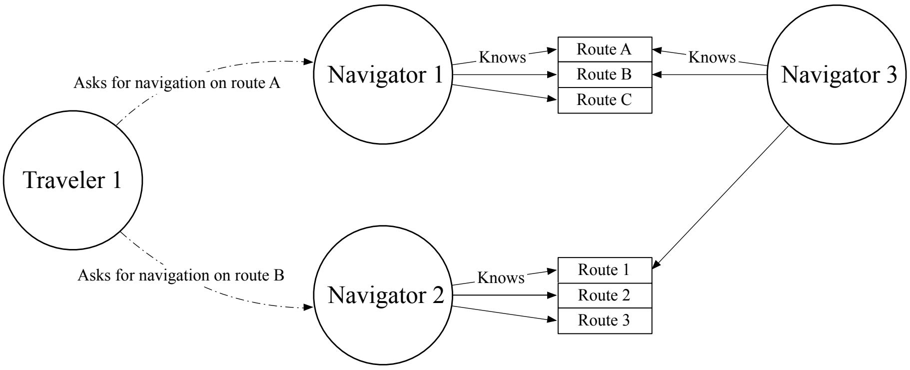


Figure 5.4: Instance diagram of the collaborative navigation system: knowledge of routes is depicted by a solid lines and interactions between users are depicted by a dash-dotted lines.

The second is to establish of a connection between the *traveler* and the *navigator*. We will discuss the most feasible solutions for both problems.

The collection of regular routes can be fully automated. The location of the *explorer* is collected only when s/he is walking (we are not interested in routes traveled by public transport or by car). The biggest problem with this solution is to recognize that the *explorer* has learned the route and can become a *navigator* on this route. One strategy can be to observe the frequency and the total number of walks on this route. We can assume that a route traveled 3 times a week for one month should be learned sufficiently. However, there is no support for such an assumption. Another strategy can be to collect regular routes via interviews. During the interview, the *explorer* describes his/her new regular routes in as much detail as possible. Subsequently, the description is plotted into the database. In this way, we can verify that the *explorer* has learned the route sufficiently to navigate other people [8].

In order to connect the *traveler* and the *navigator*, we need the approximate location of the *traveler* to match his/her position with available regular routes of all *navigators*. The location of the *traveler* is sent to the system, and then the system can hand the telephone numbers of suitable *navigators* over to him/her. The *traveler* can then select a *navigator* from the list, or use the first name recommended by the system.

We have observed the course of dialogs for localizing the *traveler* precisely in various types of environments (indoors, in an outdoor open space, in an outdoor urban area). All experiments shared a similar dialog model, suggesting that the key to successful navigation is to find an unambiguous orientation point in the vicinity of the *traveler* [109], [9].

5.3.5 Game Aspects of Collaborative Navigation

The collaborative aspect of the navigation application can be used to introduce game elements. As the density of regularly walked routes is essential for the quality of the navigation service , it is desirable to stimulate exploration of new routes, willingness to help others, and the courage to travel in unknown areas.

To introduce game elements successfully into our system, we need to define our system according to the definition of a game chosen in the previous section [90]. Our system can be defined as an open experiential game. The objects in the game are three players: *traveler*, *navigator* and *explorer*. The attributes of each player are the regular routes that they know, their current status (e.g., rank, badge – see below), their current position in the environment, their availability to give assistance, etc. The internal relationships between objects (players) include interaction with the environment that they are currently located in and communication with each other. The environment is an urban area (typically a city) where the players move, and the navigation aids that they use. The artificial conflict is represented by a competitive contest, e.g., two *navigators* trying to get a better score (see subsection 5.3.5), or a cooperative contest (e.g., *navigator* and *traveler* cooperating to find the desired destination). The main objective is to create a dense network of routes and a sufficient number of players to ensure uninterrupted functioning of the navigation system. Individual objectives are represented by a system of ranks, badges and points, which are described in detail below. The basic rule states that the number or the length of a route is counted when the user acts in one of the roles (*traveler*, *navigator*, *explorer*). There are also other rules that define how the player can gain a higher rank on the basis of the quality of his/her behavior (e.g., high quality of the navigation instructions, the difficulty of the route).

Rank

Ranks correspond to the user roles. There are ranks for *navigator*, *explorer* and *traveler*. Each rank has several levels. In our case, we use ranks similar to those used in the army. Higher ranks are awarded for various kinds of activities performed in each user role when using the collaborative navigation system. Higher ranks for a *navigator* are gained for the number of navigated *travelers* multiplied by the subjective evaluations made by the *travelers*. For *explorers*, higher ranks are awarded for the length of the new routes added into the system, multiplied by their complexity. *Travelers* are awarded a higher rank for the number of requests for navigation multiplied by the success rate (whether the destination is reached). In ascending order, the ranks are captain, major, general. The ranks are used as a prefix to the user role (e.g., Captain *Explorer*, Major *Traveler*).

Badge

Badges are awarded for specific behavior. There are badges used for various user roles, and badges that can be awarded for all users of the system. The *navigator* receives a badge for making him/herself available for navigating during the night hours (after 22:00, for being available for navigation non-stop, or for being the most active *navigator* in a certain area ("a local guru"). An *explorer* can get a badge for learning extreme routes (e.g., in a park or on a nature trail). A *traveler* can win a badge for using the service intensively. All users can get a badge for the highest number of points collected during the previous week. There is another set of badges that a user can award to him/herself. There are badges for using a dog or for the preferred type of transport (using a subway, a tram, a train, or going on foot). These badges can be used when selecting a *navigator* according to the *traveler's* preferences (e.g., if s/he wants to be navigated by someone who also prefers traveling by tram rather than using the subway).

Points

Points are used for determining the top user of the week in each user role (the winning user is awarded a Top User badge). Points for *navigators* are awarded for the number of *travelers* navigated. For *explorers*, points are awarded for the length of new routes added into the system. *Travelers* are awarded points for the number of requests they make for navigation.

The badges, ranks, and points are presented on users' profiles and can be viewed by all the users. They are tokens of high activity and enthusiasm in using the collaborative navigation system. They also reward and reinforce some high-quality performances and positive behavioral patterns of the winners. Badges, ranks, and points therefore also provide a basis for recommending, recognizing or selecting a good *navigator* when a *traveler* seeks help with navigation in a place where more than one *navigator* is available.

5.4 Conclusion

In this chapter, we have shown how we can use gamification to enhance a collaborative navigation systems for navigation of people with visual impairments. In the gamification process, we have focused on supporting the cooperative nature of the system. We have incorporated high-level game objectives like building up a dense network of routes, as well as individual objectives, i.e., ranks and badges.

An interesting question is how to attract other potential actors, e.g., friends or the general public. If they will find it interesting to play our navigation game, we can further improve the quality of the navigation system. It could be interesting to introduce new game characters and activities which will attract sighted participants into the audience.

5.4.1 Future Work

In future work, we should focus on compiling rules and strategies for successful navigation dialog among visually impaired users of the system. The further research should also focus on studying the influence of stress visually impaired people are suffering from while traveling. To achieve and preserve the critical mass of users of our collaborative navigation system, the integrated game aspects should be improved. In particular, extend the use of points gained for using the system, for example by making the points exchangeable for real benefits, such as vouchers, tickets, etc.

5.5 Implications

In this chapter, we presented a theoretical concept of a collaborative navigation system. We proposed a design enhanced with gamification elements to increase user engagement. The concept is based on findings presented in Chapters 2 and 3. For the future research, one could focus on criteria used for matching the users. Apart from services like Be My Eyes[1] which connect sighted volunteers over a camera with a visually impaired person to respond to visual questions [5], one may aim to match users based on cognitive styles and navigation strategies.

For the next research focus, we took a strategic decision to focus more on automation and substitution of a real person with a conversational agent. Thus, the next two chapters focus solely on natural language generation (Chapter 6) and dialog modeling and design (Chapter 7) for navigation conversational agent. We leave the implementation and evaluation of collaborative navigation open to the future work.

---

[1] Be My Eyes – https://www.bemyeyes.com/

Chapter 6

Automatically Generated Verbal Description

Abstract. *Visual impairment limits a person mainly in the ability to move freely and independently. Even with many navigation aids and tools currently on the market, almost one-third of the visually impaired do not travel alone without a guide. One of the most useful aids are the landmark-enhanced itineraries of the route created by orientation and mobility specialists. We describe the design of a system, which is based on a specific, efficiently collected geographical data, and which automatically generates human-like landmark-enhanced itineraries for navigation of visually impaired users. The studies we conducted (quantitative* $n = 16$*, qualitative* $n = 6$*) showed usability and efficiency of the system. We also provide a set of design recommendations to increase the usability of the system along with specific examples of usage of particular landmarks. Moreover, we report on 14-day long diary study (*$n = 3$*) where visually impaired participants used the system freely when walking throughout the city center with a success rate of 84 % without using GPS. Finally, based on an over-the-phone survey (*$n = 18$*), we provide a methodology to identify the essential area, where geographical data needs to be collected to enable visually impaired pedestrians to travel around the city center efficiently.*

Note. *This chapter is based on published journal paper Landmark-enhanced Route Itineraries for Navigation of Blind Pedestrians in Urban Environment [14].*

6.1 Introduction

In the area of electronic navigation aids and systems, the problem of the visually impaired persons' mobility is currently solved by means of car navigation systems, which are not suitable for the visually impaired. The itinerary generated by such a car navigation system lacks important information about the sidewalks, landmarks, crossings, etc., which is essential for safe and efficient traveling (see Subsection 6.2). Another approach is the use of an assistance of orientation and mobility specialists [104]. These specialists can prepare a description of a route to a remote destination for a visually impaired person for a particular route in advance and provide it in the form of a route itinerary. This solution suffers from time requirements, as it has to be prepared in advance, and rigidity, as there is no option to change the route at user's will.

In this chapter, we focus on designing a concept of an efficient navigation aid which supports navigation to a remote destination. We aim to generate route itineraries similar to those prepared by orientation and mobility specialist [4]. By using sophisticated data structures and algorithms, we address the issue of time requirements (the itinerary is created instantly), rigidity (user can select whichever origin and destination s/he wants) and dependency on other people (passersby or orientation and mobility specialist).

To compare our solution to state-of-art electronic navigation systems, we designed two versions of the navigation system. The first version (*Landmark*) was prototyped with itineraries enhanced by landmarks. The second version (*Metric*) was prototyped to simulate current metric-based navigation systems (i.e., “After 90 meters turn right” [85]). For the fair comparison of both versions, the *Metric* one also used a pedestrian network for routing (sidewalks, crossings). Next, we provide the insights from the qualitative study conducted on the *Landmark* version to further investigate the role of the landmarks. Further, we report on a 14-day long diary study we carried out to observe long-term usage of the revised *Landmark* version by visually impaired pedestrians in the wild. Finally, we provide a method and its results to identify the area, which needs to be covered by the future navigation system to the meet needs of the visually impaired persons visiting the city center.

6.2 Related Work

In this section we provide additional related work on navigation aids for visually impaired. The related work on verbal description its relation with cognitive maps can be found in Chapter 1.

6.2.1 Navigation Aids for the Blind

There are numerous navigation aids for visually impaired pedestrians. Some use special sensors to identify objects on the route, e.g., cameras [23], or an RFID based electronic cane [39]. Others are based on a concept described in [80], and rely on some kind of positioning system (e.g., GPS) in combination with the GIS system to identify objects and navigate the pedestrian with visual impairments, e.g., Ariadne GPS, BlindSquare. There have also been attempts to develop special interaction techniques for presenting navigation instructions, e.g., an auditory display [62] or a tactile compass [81].

The navigation aids based on major GIS systems (such as Google Maps, Apple Maps, OSM Maps, or Nokia HERE Maps) suffer from an inappropriate description of the environment for visually impaired pedestrians. The available description may be imprecise (e.g., missing sidewalks or missing handrails), or may be ambiguous (e.g., an inadequate description of pedestrian crossing, meaning that it cannot be localized and identified without visual feedback) or it may ignore specific navigation cues (e.g., the surface structure of the sidewalk, acoustic landmarks such as the specific sound of a passage, the traffic noise of a busy street, or other sensory landmarks, such as the smell of a bakery). In addition, routing algorithms can encounter problems with non-trivial adjustments to the preferences and abilities of visually impaired people, e.g., their inability to cross open spaces (e.g., large squares). Moreover, visually impaired people report imprecise GPS signal and lack of the data which can be used by GPS navigation[116], that is because of low accuracy of GPS in cities – 28 meters inaccuracy 95 % of time [71].

Both inappropriate descriptions and unsuitable routing algorithms can be avoided by introducing navigation systems based on tele-assistance with a trained human agent. Various approaches have been proposed on the basis of various ways to identify the position and the environment of the pedestrian, like transmission of chest mounted camera view to the navigator bujacz2008, a verbal description from the pedestrian optionally combined with GPS location and GIS [30], [100], or purely based on a verbal description and knowledge of the environment [109]. Namely Navigational Centre for the Blind sons, operating since 2007, proved to be helpful tele-assistance navigation service widely used (6650 cases in years 2008–2013 [29]) by community of visually impaired people.

To support independent mobility of visually impaired travelers we aim to generate spatial language which inherently contains landmark information [32]. We include important relation between landmarks to support efficient landmark to landmark navigation [47], [67], [88] and actions to perform related to egocentric frame [68], [69]. Moreover, the use of spatial language supports creation of mental representation of a route [75], updating spatial images [64] as opposed to “car-like” automatic navigation systems providing metric instructions and disengaging user from navigation process [61] and acquisitions of spatial knowledge [76].

In this work, we consider an amodal spatial language as a modality substitution for vision used to create mental representation of a route and to update spatial images. The user interface of the system (in the last version – see Section 6.5) outputs amodal spatial language using auditory output modality (text-to-speech), uses tactile input as an interaction method (gestures), and auditory modality as an input (voice).

6.3 Landmark-enhanced Route Itineraries Generation

To implement a feasible solution to generate landmark-enhanced route itineraries, we needed a specially modified GIS capable of representing special features of the urban environment. Furthermore, we also needed 1) algorithms which use the landmarks and their parameters to generate route itinerary in a natural language and algorithms, and 2) algorithms which use landmark parameters to find an optimal route from the origin to the desired destination.

In our case, each route itinerary is composed of navigation instructions for each segment of a route divided by decision points (typically a part of a route from a corner to a corner, from a corner to an entry to a crossing, or from an entry to a crossing to the other side of the street). For each navigation instruction, we have chosen the following structure: environment description and action that should be performed by the blind pedestrian.

6.3.1 Geographical Information Database

To implement a feasible solution to generate landmark-enhanced route itineraries, we needed a specially modified GIS capable of representing special features of the urban environment.

We extend a regular street link network used for car navigation with new pedestrian links based on sidewalks. We distinguish line, point, and area features:

- **Line features** are tied to large part of the pedestrian link (e.g., geometry representation of a sidewalk in GIS) and represent their properties (like slope, surface quality) or phenomena along the link (e.g., parking cars, railings).

- **Point features** describe phenomena that cover a small part of the pedestrian link (typically less than 3 m), they act as barriers or landmarks (e.g., crossing entry points, recesses, steps, corners).
- **Area features** are landmarks extending over a certain area. Those are phenomena such as traffic noise (e.g., busy streets or certain types of public transport).

The nature of those new features and the range of descriptive information reflect the needs of visually impaired people. The requirements are based on findings of Balata et al. [9] who observed how pairs of visually impaired people navigated each other over the phone (following on research of Bradley and Dunlop [20]), including landmarks and orientation strategies leading to successful navigation (e.g., using building and traffic noise as orientations cues at beginning of a route). Further, the requirements are based on Volkel et al. [107] survey where they explored mobility of people with visual impairments from data requirements and routing perspective (e.g., avoiding large plazas). However, our GIS can be universally used for navigation of other groups of people with disabilities such as wheelchair users or older adults. The complete data model was designed in cooperation with Central European Data Agency, a.s. (CEDA), Faculty of Transportation Science and Faculty of Electrical Engineering (Czech Technical University in Prague) within project ROUTE4ALL. The data model was implemented as an ArcGIS Geodatabase[1].

Pedestrian link network

To be able to collect desired landmarks and their properties, the existence of the pedestrian link network is essential (similarly to Volkel [108]). Thus, the national provider of digital vector road geodatabase StreetNet created a new product PedestriNet that represents all paths designated for pedestrians. StreetNet is currently used for pedestrian navigation for users without disabilities. For people with disabilities, it is not suitable due to the simplified representation of crossings and sidewalks that does not reflect the real topology of the pedestrian network.

PedestriNet is characterized by a very high positional accuracy - it represents footpaths in their real position in the level of “Map of town utilities” (see Figure 6.1) and maintains the topology of the pedestrian network. It covers special pedestrian links like sidewalk, crossing, square link (a sidewalk crossing large open area of a square) as well as pedestrian links already represented in the StreetNet database (e.g., pedestrian zone, walkway, gallery).

Geodatabase

All features of geodatabase include reference information to PedestriNet – ID of the links, orientation towards it (left/right/on). Thus they can be used for routing along the pedestrian network and generation of route itinerary.

To minimize the data collection costs, it is necessary to use existing data as much as possible. City municipalities usually administer “Map of town utilities” in high accuracy level that may be used to locate some features (e.g., stairways, corners) or to derive some properties (e.g., width

---

[1] ArcGIS – www.arcgis.com/

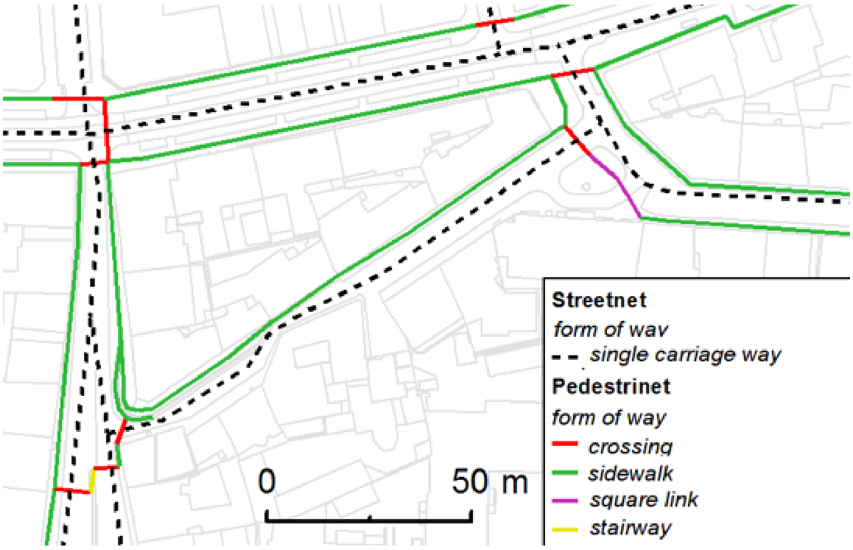


Figure 6.1: Comparison of StreetNet and PedestriNet – background data: Digital city map - Map of town utilities (source: Open data).

of a sidewalk). Unfortunately, these maps are mostly Computer Aided Design (CAD) drawings, and therefore full automation of data processing is impossible. Further, a "Land-use map" can be used to derive land-use of adjacent footpath area; a digital terrain model can be used to calculate a slope of a footpath, etc. Another source of information may be a spherical video recording (i.e., Google Street View). Howerver, the field survey remains necessary to derive precise data such as a height of curbs and to verify data obtained using remote methods (e.g., using CAD drawings, satellite imaging). Here, a community reporting and crowdsourcing may be of high importance to eliminate errors and keep data up-to-date [86].

A prototype of the geodatabase that is used for validation and testing covers the area of about 0.8 km^2 in the city center which in this case corresponds to 51.9 km of pedestrian links. The area was chosen to both capture representative example of an urban environment and to be accessible from our university for the experiments. Creation of this area took around five man/days to collect data and fill the geodatabase. However, it is not necessary to collect data for the area of the entire city (see Section 6.7).

6.3.2 Optimal Route Finding

Finding a suitable route in a city is challenging task as the shortest route is not always the most appropriate for blind pedestrians. The algorithm has to take into account specific needs and abilities of the blind pedestrians such as preference to use pedestrian crossings with acoustic signalization, avoidance of large open spaces, etc. The geodatabase contains useful information about pedestrian crossings, stairs, passages, sidewalks without a curb, or paths in parks. This information is used to determine a penalty, which is then assigned to a corresponding pedestrian link and it is used in the algorithm (see paragraph Routing algorithm).

Geodatabase pedestrian link parameters utilized to determine the penalty are the following:

- the average intensity of traffic on the street,
- the length of the passage,

- the presence of the curb,
- the avoidance of stairs,
- the traffic light and sound signalization on the pedestrian crossing,
- the "findability" of the pedestrian crossing[2].

The introduction of penalties enabled us to apply standard Dijkstra algorithm to find the shortest route from an origin address to a destination address. The penalties are optimized to prefer safer route even at a the cost of longer distance.

Routing algorithm

We implemented a Dijkstra's algorithm variation with binary heap as a priority queue. We have compared implementations of a binary heap with and without decrease-key operation. However, the one without decrease key performed better on the graphs generated from the geodatabase. We also implemented associative arrays using hash tables to keep track of visited nodes and predecessors to avoid storing this information within graph nodes. This enabled us to store algorithm information within nodes without need to re-initialize all nodes before each algorithm run. Since the graphs generated from the geodatabase are huge, this could lead to a performance problem.

To support penalizations, the length of graph edges is adjusted based on their associated pedestrian link parameters before running the algorithm. The edges have an associated function to calculate their length (adjusted for penalization) in runtime. These functions can be arbitrary code. We implemented several sets of functions (and their parameters), which are executed for each route search, to select the best route.

The routing algorithm is based on parameters which can be changed and can influence the routing. However, for purpose of the studies, we did not let the participant select from more routes or adjust the parameters. In the experiments we focused namely on automatically generated route itineraries – letting users select from multiple options or adjust the parameters is a matter of future work.

6.3.3 Navigation Instruction Structure

As described earlier (at the beginning of this section), each navigation instruction is composed of environment description and action that should be performed by the blind pedestrian (similarly to [110] – description + action and in line with [32] – references to landmarks and actions in spatial language). The environment description is generated from street names, addresses, corners, and crossings. The action is generated from geometry (pedestrian links), street names, corners, slopes, land-use, and point features. Table 6.1 shows how each navigation instruction is composed on a short five segment route itinerary with one crossing.

---

[2] A pedestrian crossings located near corners are easier to find than those in the middle of the block of buildings

Listing 6.1: Selecting best instruction template.

```
ISituation actualSituation = new StartSituation();
ISituation endSituation = new EndSituation();

while(actualSituation != endSituaton) {
  ITemplate template = FindBestTemplate(actualSituation);
  ISituation actualSituation = template.Apply(actualSituation);
}

function FindBestTemplate(ISituation situation) {
  ITemplate[] availableTemplates = GetAvailableTemples();
  ITemplate bestTemplate = availableTemples
        .Where(template => template.Accepts(situation))
        .OrderByDescending(template => template.Priority);
  return bestTemplate;
}
```

Listing 6.2: Application of an instruction template.

```
class CornerDescriptionTemplate : ITemplate {

  function Accepts(situation) {
    return IsAtCorner(situation) && situation.IsNotApplied(this);
  }

  function Apply(situation) {
    templateText = “You are at a~{0} corner of streets {1} and {2}”;
    FillSituationVariables(template);
    return SituationWithAppliedTemplate(this);
  }

}
```

Template algorithm

First, a route is found by an algorithm on PedestriNet pedestrian link network (see Section 6.3.2). Then, for each segment (formed of one or more pedestrian links), which is represented by vector data, we create a navigation instruction.

To generate navigation instruction we find best matching instruction templates, which are selected based on a type of action that should be performed (turning, crossing a street, etc.), context (available metadata, adjacent segments, direction of the user, etc.), and priority (e.g., the template for corner is more preferred than the template for crossing) (see Listing 6.1). For example, result of an instruction template for environment description of a place is: *“You are at a beveled corner of streets Odboru and Karlovo namesti.”* (see Listing 6.2).

In this way, we created landmark-enhanced navigation instruction and optimal routing for blind pedestrians (*Landmark* version), which were later compared with metric-based navigation instructions (*Metric* version).

No.	**Env. description**	**Action**						
x/y	*Startpoint*	*Direction*	*Motion*	*Distance*	*Slope*	*Endpoint*	*Properties*	*Land-use*
1/5	You are at the address Karlovo námesti 293/13.	Turn to the left	and go	approximately 150 meters		to the round corner with Odboru street.		Keep the buildings on your left-hand side.
2/5	You are at the round corner of Karlovo namesti and Odboru streets.	Continue straight	and cross Odboru street			to the opposite corner	via crossing with light signalization and one-way traffic from right.	
3/5	You are at the beveled corner of Karlovo namesti and Odboru streets.	Turn to the left	and go	approximately 100 meters	slightly downhill	to the corner with Myslikova street.	The street bends to the right.	Keep the buildings on your right-hand side.
4/5	You are at the corner of Odboru and Myslikova streets.	Turn right	and go	approximately 30 meters		to the address Myslikova 282/26	which will be on your right-hand side.	Keep the buildings on your right-hand side.
5/5	You are at the destination. You are at the address Myslikova 282/26.							

Table 6.1: Main building blocks used for automated generation of route itineraries (translated from Czech). The itinerary is composed of segments (rows), formed of environment description (second column) and action to be performed (third column).

6.4 Quantitative Evaluation

In this experiment, we raised the question whether the error rate is lower for *Landmark* version than for *Metric* version and whether measured completion time is lower for *Landmark* condition than for *Metric* condition.

Further, we investigated subjective judgment of the participants about the level of safety, comprehension, and ambiguity of the generated itineraries, along with qualitative observations.

6.4.1 Participants

Sixteen visually impaired participants (10 female, 6 male) were recruited via e-mail leaflet sent to a pool of long-term collaborators of our university. The participants in the experiment were aged from 23 to 66 years ($mean = 35.75$, $SD = 11.23$). Eleven participants had Category 5 visual impairment (no light perception); 5 participants had Category 4 visual impairment (light perception) [113]; 4 participants were late blind, 12 participants were congenitally blind. All of the participants were native Czech speakers (see Table 6.2). The participants can be considered skilled travelers – they arrived alone to the university and they clearly do not belong to 30 % of visually impaired people who never travel alone [28], [112].

Participant	**P01**	**P02**	**P06**	**P07**	**P08**	**P09**	**P10**	**P11**	**P13**	**P14**	**P15**	**P16**	**P18**	**P19**	**P20**	**P21**
Gender	F	F	F	F	F	F	F	M	M	M	F	F	M	F	M	M
Onset	L	L	C	C	C	C	C	C	C	C	C	C	C	L	C	L
Category	4	4	4	4	5	5	5	5	5	5	5	5	4	5	5	5
Age	66	40	39	30	23	29	39	44	30	34	26	27	30	24	51	40

Table 6.2: List of the participants in the comparative study, with age, the onset of the impairment (C – congenital, L – late) and the category (4 – with light perception, 5 – without light perception) of the impairment (participant IDs are not sequential due to recruitment process).

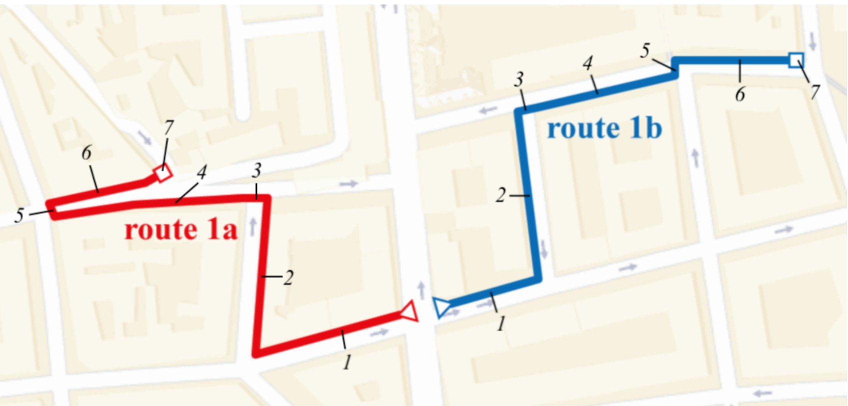


Figure 6.2: The routes used in the comparative study; the triangle depicts the beginning of a route; the square depicts the destination of a route; numbers represent the segment's numbers.

6.4.2 Apparatus

Routes

For our within subject experiment, each participant had to be exposed to both *Landmark* and *Metric* based route itineraries. Similarly to [57] we have chosen two routes that can be considered as equivalent (same length, same number of decision points, same number of segments, etc.). By choosing two equivalent routes we eliminated the learning effect on one hand and minimized the influence of differences of the routes on the experiment on second hand. Environments for this type of experiment are usually real environments [20], [85] rather than artificial (lab) environments, though exceptions are possible [93]. Thus in our experiment the two routes were situated in a quiet area in the city center of Prague, Czech Republic (see Figure 6.2). Both of the routes were approximately 350 meters long, consisted of 7 segments and 8 decision points (points where the participant changes his/her direction of movement). They had the same number of turns and pedestrian crossings.

The selection of the routes aim to contain typical situation encountered when travelling in the city – walking along building, along park, crossing the street (different kinds, lengths, traffic), walking through passages, etc. In this way we aim to observe behaviour applicable all routes in given context (central European city – culturally similar region with similar city layouts).

Equipment

The participant was equipped with a Nokia 6120 mobile phone with a lanyard which hung from his/her neck. In this way, the phone was protected from being dropped unintentionally, and the

No.	***Landmark* version**	***Metric* version**
1/7	You are at the corner of Resslova and Dittrichova street. Continue straight and go approximately 60 meters to the corner with Jenstejnska streets. Keep the buildings on your left-hand side.	Continue straight and go 55 meters on Dittrichova street.
2/7	You are at the corner of Dittrichova and Jenstejnska streets. Turn left and go approximately 90 meters slightly uphill to the corner with Vaclavska street. Keep the buildings on your left-hand side.	Turn left and go 85 meters on Jenstejnska street.
3/7	You are at the corner of Jenstejnska and Vaclavska streets. Turn to the right and cross Jenstejnska street via unmarked crossing with one-way traffic from left.	Turn to the left and cross the street.
4/7	You are at the corner of Jenstejnska and Vaclavska street. Continue straight and go approximately 70 meters to the corner with Trojanova street. Keep the buildings on you right-hand side.	Continue straight and go 74 meters on Vaclavska street.
5/7	You are at the corner of Vaclavska and Trojanova streets. Turn to the left and cross Vaclavska street via crossing with one-way traffic from right.	Turn to the left and cross the street.
6/7	You are on Vaclavska street. Turn to the right and go approximately 80 meters to the corner with Na Morani street. Keep the buildings on your left-hand side.	Turn to the right and go 75 meters on Vaclavska street.
7/7	You are at the corner of Vaclavska and Na Morani streets. Turn to the left and go approximately 1 meter to the destination Cafe Amandine. Keep the buildings on your left-hand side.	Turn to the left and go 1 meter on Na Morani street. Destination Cafe Amandine will be on the left.

Table 6.3: Route itineraries of *Landmark* version and of *Metric* version used in quantitative evaluation (see Section 6.4) for route 1b (see Figure 6.2), route 1a was created in a similar manner.) (translated from Czech).

participant was able to release it and have an empty hand when needed, and s/he could also grasp it again quickly. The mobile phone was set to the Czech language, and it was equipped with the MobileSpeak text-to-speech (TTS) screen reader application by CodeFactory. A navigation application prototype was implemented in Java ME and it used text files with pre-calculated route descriptions for corresponding routes (see Table 6.3). The participants interacted with the navigation application prototype by means of pressing buttons for “next segment”, “previous segment”and “repeat current segment” mapped on functions keys and one menu item (previous segment).

Data collection

In each session, we recorded two video streams of the participant’s activities. The first camera (GoPro Hero 3) recorded 1st person view and was installed on a shoulder strap of the backpack that was carried by the participants during the session, while the second camera (Sony DLSR) recorded a 3rd person view by the experimenter shadowing the participant.

6.4.3 Procedure

The experiment consisted of two walkthroughs of each route and it lasted around 1.5 hours. In the first walkthrough the experimenter guided the participant to the beginning of the first route, explained the purpose of the experiment to the participant, explained operation of the navigation application prototype, and asked the participant to adjust the phone on a lanyard or to hold it in a hand, according to his/her own preference. The participant was asked to proceed as quickly and accurately as possible. The task was given as follows: *"You have a meeting in Hostel Emma (for route A; Cafe Amandine for route B). To reach the destination use the navigation application prototype. Proceed as if you were alone, but we will be watching for your safety from a distance."* Then the participant started out the first walkthrough.

If the participant went astray (made navigation error) the experimenter intervened in the experiment when one of the following conditions were met — when the participant started crossing on other place than on pedestrian crossings; when the participant went of the planned route for 5 meters.

After the first walkthrough, the participant was returned to the start of the route and walked the route with the experimenter. The participant was retrospectively asked about his/her subjective judgment about the level of safety ("I felt safe on this segment."), comprehension ("I knew exactly how to further proceed.") and ambiguity ("There was nothing in the environment that confused me?") for each segment of a route. Likert scale 1-5 was used, i.e., expressing level of agreement with presented statement – strongly agree, agree, neutral, disagree, strongly disagree.

Then the experimenter took the participant to the start of the second route and proceeded same as on the first route. After the experiment, the participant was debriefed and received his/her payment.

6.4.4 Design

The experiment was one factor (two levels) within subject design. The independent variable was itinerary version (*Landmark*, *Metric*). The order of itinerary version and routes was counterbalanced using a Latin square. The main measures were an error rate, defined as a ratio of number of segments with navigation error to total number of segments (as a navigation error we consider a situation when participant missed the turn, crossed the road in a different place, turned and walked in a wrong direction), and a completion time, calculated as time elapsed from start of the route to the destination of a route (including stops for navigation instruction listening or time spent in error situations). For analysis of error rate and completion time, we used confidence intervals (according to [37]).

6.4.5 Results and Discussion

The following paragraphs describe findings observed during the experiment. We collected the data from 16 sessions and based on the results we propose general design recommendations for the improvement of navigation instructions.

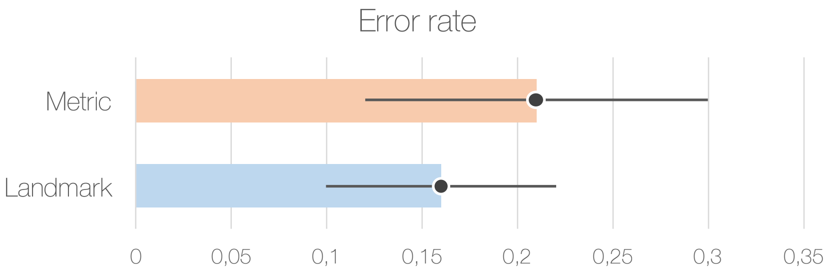


Figure 6.3: Mean error rates with 95 % confidence intervals for Landmark and Metric condition (n = 16, lower is better).

Error rate

Figure 6.3 provides evidence on error rates and 95 % confidence intervals. It seems that error rate for *Landmark* ($mean = 0.16$, 95 % CI $[0.10, 0.22]$) was very similar as for *Metric* ($mean = 0.21$, 95 % CI $[0.12, 0.30]$) and the results are largely inconclusive concerning the difference between the test conditions, although with small favor for *Landmark*.

We could not decide whether the lack of difference in error rate was caused by random variables occurring during the experiment (see paragraph Random variables) or by selection of rather easy routes as described by one participant.

Completion time

Figure 6.4 provides evidence on completion times in seconds and 95 % confidence intervals. It seem that completion time for *Landmark* ($mean = 615.4$ $seconds$, 95 % CI $[520.89, 709.99]$) is 1.2× higher on average than for *Metric* ($mean = 513.00$ $seconds$, 95 % CI $[450.33,\ 575.67]$) and the results show that there is an effect of the test conditions on completion time. The completion times ranged from 267 to 917 seconds for *Landmark* test condition, and they ranged from 292 to 757 seconds for *Metric* test condition.

It cannot be decided whether the difference in completion times for *Landmark* and *Metric* was caused by longer text in the navigation instructions (participant waited a longer time to listen to it whole) or by the occurrence of random variables (sometimes participants stopped because they did not feel secure, see paragraph Random variables). It would be necessary to repeat the experiment in more controlled conditions. However, it would affect the external validity of the experiment. Similar problem was also found by [85] whose completion times were affected by red-light waiting times at crossroads.

Subjective judgment

During the second walkthrough, we asked the participants about their subjective judgment of each segment of a route. The results suggest that comprehension is higher for *Landmark* (85 % of the participants strongly agree) than for *Metric* (65 % of the participants strongly agree). Ambiguity and safety were evaluated similarly for both test conditions (see Figure 6.5).

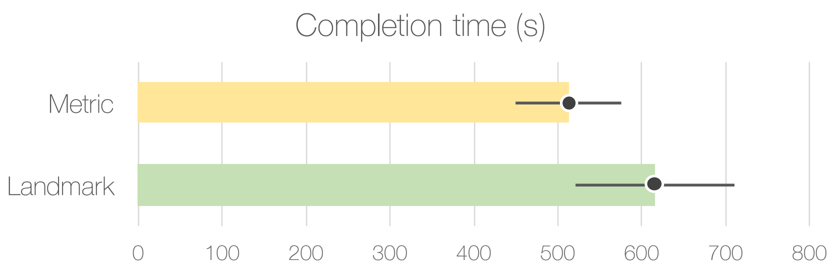


Figure 6.4: Mean completion times with 95 % confidence intervals for Landmark and Metric condition (n = 16, lower is better).

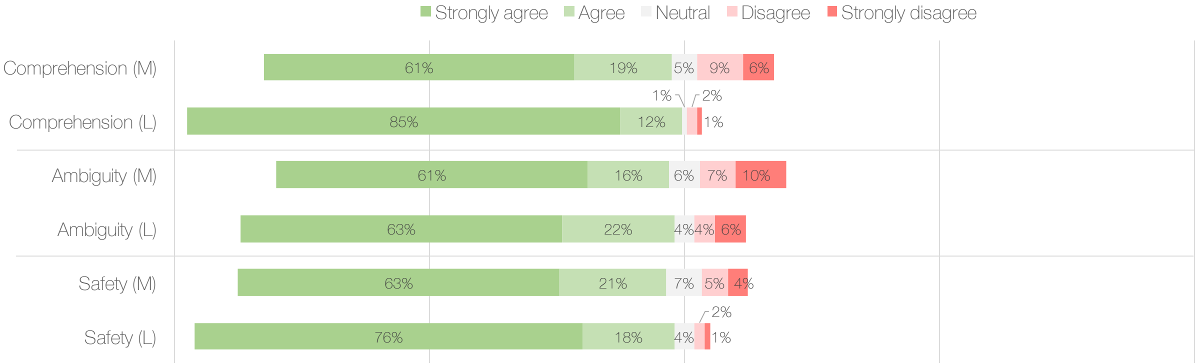


Figure 6.5: Subjective judgments about the level of safety, comprehension, and ambiguity of navigation instructions, summarized over all segments, for both conditions – (L) *Landmark*, (M) *Metric* (n = 16).

6.4.6 General observations

Further, we asked the participants for the comments on navigation instructions for both conditions.

On *Landmark* condition, the participants highlighted usefulness of information about traffic direction or information about land-use. However, nine participants had problems with lack of the corner descriptions – they had troubles orienting themselves on the corners with different shape than 90 degrees, i.e., when asked on round corner to turn to the right, they often proceeded in wrong direction. Further, the participants missed information about endpoints at crossings – they did not know, what to expect on the other side of the street, whether there is a sidewalk or corner of a building. On *Metric* condition, the participants missed information about endpoints and they were surprised by very precise distance indications (precision up to single meters; in *Landmark* condition, distances are rounded to tens of meters) reported by the system (*"I cannot tell how far did I go precisely,"* P07, participant P01 laughed about precise distances – she found the precision funny as she can never guess the precise distance). As the description lacked information about endpoints (landmarks at the end of segment), the participant were often confused and did not know how far should they go. They solved this by listening to the next navigation instruction to know how will they proceed next.

On both test conditions, the participants often interchanged navigation instructions of succeeding segments as all navigation instruction started similarly – *"You are at the corner of..."*

– (the navigation instructions in tested version lacked the numbering as it can be seen in Table 6.1 first column).

One participant mentioned that s/he would benefit from the usage of GPS geofencing, which would notify him/her about next navigation instruction. Two participants found the *Landmark* version too detailed and they would prefer *Metric* version on routes they already know (we address this request by means of conversational navigation agent, see Chapter 7).

Random variables

Similarly to Rehrl et al. [85], our measurements were influenced by random variables (unpredictable urban environment). Further, we classify the problems we observed during the experiment into four categories.

Collision with objects – 27× We observed that participants frequently collide with traffic signs, poles, beer gardens or parking ticket machines. This seems like a common problem for the visually impaired pedestrians.

Interference of passerby people – 11× During the experiment, we observed the participant from a distance (shadowing method). Sometimes passerby people stopped the participants and offered them help (5×), however, 6× they grabbed, dragged or guided the participant to some arbitrarily chosen spot, which resulted in a loss of orientation of the participant.

Disruption of the senses – 17× As a hearing is one of primary orientation and navigation senses for visually impaired pedestrians, its disruption strongly affects the wayfinding process. We observed following sources of hearing disruption: garbage disposal trucks, road cleaning trucks, and rain (cars driving on a wet roadway produce much louder noise). Another case was a problem with finger sensitivity and operation of our Nokia smartphone of one participant in both sessions.

Stress – 3× One participant was anxious about the experiment even though we tried to calm him/her down during the briefing and s/he appeared calm before start of the experiment, the stress affected his/her performance on both conditions (s/he proceeded much better towards the end of the experiment than in the beginning). Another case was when a participant dropped his/her white cane (s/he was immediately assisted and handed the cane).

Altogether, we counted 31 common problems (collisions and offering a help of passersby people), 16 in *Metric* conditions and 15 in *Landmark*. Next, we counted 27 serious problems (grabbing, dragging, guiding by passerby people, disruption of senses, stress), 16 in *Metric* condition, 11 in *Landmark* condition.

For more extensive analysis of random variables see Appendix: Random Variables.

6.4.7 Recommendations for Design

Following design recommendations for creation of navigation instructions were extracted from the findings collected during the experiment:

- **R1** The shape of a corner should be added to the itinerary if the shape is different from sharp/plain.

  We observed that participants were confused at corners, which were not sharp but beveled or rounded.

- **R2** The endpoint should be added for pedestrian crossings like the sidewalk on the other side of the street, the opposite corner, the opposite side of the street.

  The participants expected the information about the other side of the street when they used a pedestrian crossing.

- **R3** Change the beginning of the navigation instructions by adding a sequential number and a total number of navigation instructions.

  Some participants were confused when the beginning of succeeding navigation instructions was the same (i.e., *"You are at the corner of Jenstejnska and Vaclavska streets..."*. This occurs in situation when the participant was crossing the street from one corner to the other one).

- **R4** Changing the naming conventions for pedestrian crossings without *"zebra"*.

  One of the participants expressed concerns about the usage of term *"unmarked crossing"* instead of *"a place for crossing"* for a place where there are lowered curbs but no "zebra" painted on a street.

6.5 Qualitative Evaluation

After the first experiment, we implemented recommendations R1-R4 into the algorithms and geodatabase of our system. We further investigated automatically generated itineraries in a different, much more complicated urban environment (busy streets, park, passages).

6.5.1 Participants

Six visually impaired participants (3 female, 3 male) were recruited via e-mail leaflet sent to a pool of long-term collaborators of our university. The participants in the experiment were aged from 30 to 68 years ($mean$ = 48.67, SD = 15.74). Three participants had Category 5 visual impairment (no light perception); 3 participants had Category 4 visual impairment (light perception); 4 participants were late blind, 2 participants were congenitally blind. All of the participants were native Czech speakers. They can be considered skilled travelers – they arrived alone to the university. Originally we recruited 7 participants however participant P02 canceled the appointment. Participants P01, P03, P04, and P07 participated also in previous study (see Table 6.4).

6.5.2 Apparatus

Route

For our experiment, we selected a route in a city center outdoor environment. The route went through a busy square in a city center of Prague and ended in a quiet area (see Figure 6.6). The

Participant	**P01**	**P06**	**P04**	**P05**	**P03**	**P07**
Gender	F	M	F	F	M	M
Onset	L	L	C	L	L	C
Category	4	4	4	5	5	5
Age	66	37	30	68	40	51

Table 6.4: List of the participants in the qualitative study, with age, the onset of the impairment (C – congenital, L – late) and the category (4 – with light perception, 5 – without light perception) of the impairment (participant IDs are not sequential due to recruitment process).

route was 670 meters long and consisted of 12 segments and 13 decision points. There were 4 pedestrian crossings on a route.

Equipment and Data collection

The equipment and the data collection were the same as in the first quantative evaluation (see Section 6.4).

6.5.3 Procedure

The experiment consisted of one walkthrough of the route, and the whole session lasted about 45 minutes. At the beginning, the experimenter guided the participant to the beginning of the route, explained the purpose of the experiment and explained the operation of the navigation application prototype. The participant was asked to use think-aloud protocol. The task given to the participant was: *"You are standing in front of your house on an address Na Zborenci 276/14. To reach the destination Mibi's restaurant (a green square on Figure 6.6), where you are meeting a friend, use the navigation application prototype. Proceed as if you were alone, but we will be watching you for your safety from a distance, we will also assist you on pedestrian crossings if needed."* Then the participant started out.

When the participant reached the destination s/he was asked about his/her subjective judgment about the level of comfort ("Was the navigation instruction comfortable for the participant?"), efficiency ("Did the participant think s/he proceeded efficiently?") and safety ("Did the participant feel safe?") (Likert scale 1-5 was used). The factors of subjective judgment were selected differently from the Comparative Study (see 6.4) as the subjective evaluation was done per route, not per segment. After the experiment, the participant was debriefed and received his/her payment.

6.5.4 Results and Discussion

All participants reached the destination successfully without any serious problems.

At the second decision point (beveled corner) two participants (P03, P04) missed the corner and described it as *"unclear"*. However, they found out that they missed the corner themselves after few meters.

At the second pedestrian crossing (5th segment) all participants turned right to face the pedestrian crossing even though they were not asked to. The next navigation instruction asked the participant to turn right. However, this did not confuse P07, P03, P01. On the other hand,

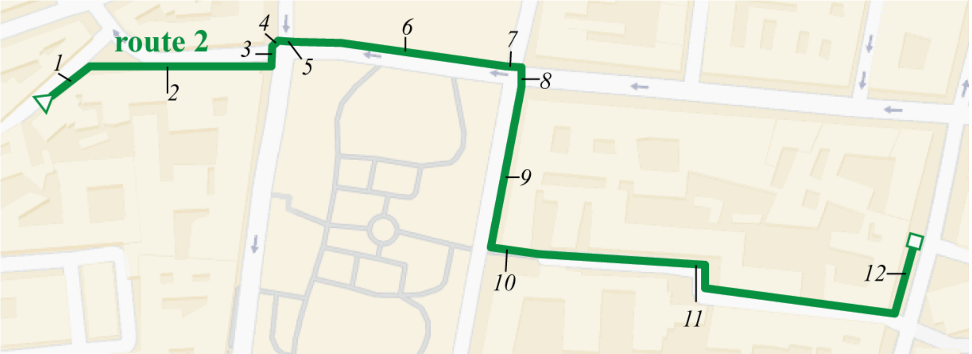


Figure 6.6: The route used in the qualitative study; the triangle depicts the origin of a route; the square depicts the destination of a route. Numbers represent the segment's numbers.

P06 and P05 crossed the pedestrian crossing immediately without selecting next navigation instruction, they did it on the other side of the street, and they got confused by the instruction to turn right and cross the street. P04 reported this as *"weird"* but was not confused by it.

At the sixth segment (sidewalk along a park) all participants followed a curb on a left side along grass even though they were asked to have the street on their right hand. P03 described that s/he used the street as an acoustic landmark on the right and followed curb on the left-hand side.

Five participants found the entrance to passage (10th segment) without any problems. The only P06 missed the passage and reported that s/he did not hear it. Participants P06 and P04 expected information about slope in a passage however the itinerary did not contain this information (it was steeply uphill).

The most problematic part of the route was 11^{th}segment where environment description contained information about the shape of the street – *"the street bends twice"*. Five participants reported that they were not sure they did perceive the second bend. P03 and P07 requested directions of the bends to be present in the itinerary. P01 did not have any problems with the bends.

Surprisingly all participants found the destination in the middle of the block of buildings (30 meters from the corner) within 5-meter precision.

Subjective judgment

After the experiment, we asked the participants about their subjective judgment about the level of efficiency, comfort, and safety during the walkthrough. Figure 6.7 shows that 33 % of the participants strongly agreed on the comfort, 50 % of the participants strongly agreed on the efficiency and 33 % of the participants strongly agreed on the safety of the navigation instructions. One participant disagreed on safety due to malfunction of acoustic signalization on a crossing.

6.5.5 Recommendations for Design

Following design recommendations were extracted from the findings collected during the experiment:

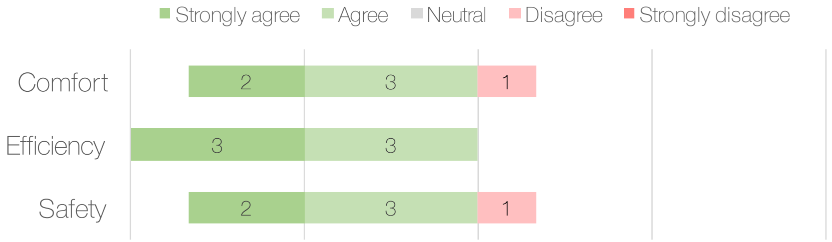


Figure 6.7: Subjective judgment about the level of safety, efficiency, and comfort of navigation instructions (n = 6).

- **R5** If the corner is more than 90 degrees use *"slightly right/left"* if the corner is less than 90 degrees use *"sharply right/left"*.

  Some participants did not recognize the corner between the first and the second segment of a route because the angle was not 90 degrees.

- **R6** If the sidewalk does not follow the building on one side, mention land-use on both sides (i.e., "there is greenery on your left-hand side, there is street on your right-hand side").

  The participants followed curb along greenery and not the street. They mentioned that they would prefer knowing both sides of the sidewalk at this segment of the route.

- **R7** If there are bends on the street mention them and add directions "first right, second left".

  The most problematic part of the route was where the street bends twice. Many participants did not perceive the second bend of the street.

- **R8** Mention sidewalk slope in passages.

  The participants expected the description of a sidewalk slope in a passage, however the navigation instruction did not contain this information.

6.6 Diary Study

We implemented the design recommendations from the qualitative study (R5-R8) into the geodatabase and algorithm generating landmark-enhanced navigation instruction (*Landmark*). To evaluate long-term effects and usage of the navigation application prototype by blind pedestrians, we conducted a 14-day long diary study (running from March 24 to April 6, 2016) with 3 participants with visual impairments.

6.6.1 Participants

Three visually impaired participants (1 female, 2 male) were recruited via e-mail leaflet sent to a pool of long-term collaborators of our university. The participants were aged from 25 to 45 years ($mean = 33.7$, $SD = 10.3$). Two participants were late blind; one participant was congenitally blind; two participants had Category 5 visual impairment (no light perception), one participant had Category 4 visual impairment (light perception) [113]. One of the participants

Participant	**P01**	**P02**	**P03**
Gender	M	F	M
Onset	C	L	L
Category	5	4	5
Age	31	25	45
Guide dog	No	Yes	No
OS	Android	Symbian	iOS

Table 6.5: List of the participants in long-term diary study, with age, the onset of the impairment (C – congenital, L – late), the category (4 – with light perception, 5 – without light perception) of the impairment, usage of a guide dog and operating system on their mobile phone.

used a guide dog. All of the participants were native Czech speakers. In the recruiting letter we stressed that we are looking for skilled travelers. None of the participants participated in two previous studies (see Table 6.5).

6.6.2 Apparatus

Participants used their own mobile phones. One participant used Android OS touchscreen smartphone, one participant used iOS touchscreen smartphone, and one participant used Symbian OS keypad smartphone.

The software consisted of accessible web navigation application prototype, which was connected to the navigation system running on a server and enabled participant to find arbitrary route (instead of pre-calculated routes in text files in two previous experiments). The web navigation application prototype was implemented and designed in agreement with WCAG 2.0[3] and WAI-ARIA 1.0[4] guidelines, and fine-tuned to be used with a screen reader software used by visually impaired users.

The main screen of the application prototype consisted of two form fields for the origin and the destination address of the route, and "search route" button (see Figure 6.8 – left). The form fields were standard HTML input elements, thus the participants were able to fill them in using standard multi-tap keyboard on Symbian OS keypad smartphone and software keyboard or dictation on Android OS and iOS touchscreen smartphones. Upon selection the "search route" button the participant were directed to navigation screen, where s/he was given navigation instruction for the particular segment of a route. The application prototype did not use GPS; thus the participants had to ask passersby for the origin address of the route or know it by heart.

The navigation screen consisted of a paragraph with navigation instruction and buttons for next and previous navigation instruction (see Figure 6.8 – right). The navigation proceeded through selecting button for the next navigation instruction after performing the current navigation instruction.

After the last navigation instruction, the participant was introduced to a screen with a questionnaire consisting of 1-5 point Likert scale (lower is better) to express his/her subjective judgment about the level of quality of both the navigation instructions and the routing, and

---

[3] WCAG – www.w3.org/WAI/intro/wcag.php

[4] WAI-ARIA – www.w3.org/WAI/intro/aria

The navigation

START AND DESTINATION

Please enter the starting address and destination address of you desired route.

Start:

Destination:

Segment 2 of 11. You are at the round corner of Karlovo náměstí and Odborů streets. Continue straight and cross Odborů street to the oposite corner via crossing with light signalization with one-way traffic from right.

Figure 6.8: A screenshot of the main screen (left) and navigation screen (right) of the accessible web application prototype used in long-term diary study.

a text area for a comment.

The interaction with the web navigation application prototype was logged on a server side. Timestamps were recorded for each interaction with the web page (searching of route, selecting next/previous navigation instruction, filling the questionnaire). Along with the timestamps, we logged the origin and the destination addresses and subjective judgment and comments filled in the questionnaire.

6.6.3 Procedure

Before the experiment, the participants were explained the purpose of the experiment and the task. The participants were asked to use the navigation application prototype as their primary navigation aid on every occasion when they had to travel through experiment area. The experiment area was selected in a Prague city center on an area of 0.8 km^2, with 51.9 km of sidewalks covered by PedestriNet geodatabase (see Figure 6.9 – green area and Section 6.3.1). The participants were asked to fill in the questionnaire about their subjective judgment about the level of quality of both the navigation instructions and the routing, and the comment after they finished each route. Additionally the participants were sent a diary on a daily basis and 5-day basis to fill in via e-mail.

The daily diaries focused on the purpose of routes they took, subjective judgment about the level of interaction comfort while using the navigation application prototype and subjective judgment about the level of assistance on the route (both on 1-5 Likert scale). The questions for daily diaries were as follows: which routes did you traveled; what was the purpose of the route; how do you access the comfort of controls while walking; how do you access level of assistance the navigation application prototype provided.

The questions for 5-day diaries consisted of those for daily diaries, and additional 3 question were added: what would you criticize about the navigation; what would you compliment about the navigation; would you recommend the navigation to your friends and family.

After the 14 days of the experiment, the participants were asked to fill in the final diary (questions about overall evaluation) and they were invited back to our university for debriefing

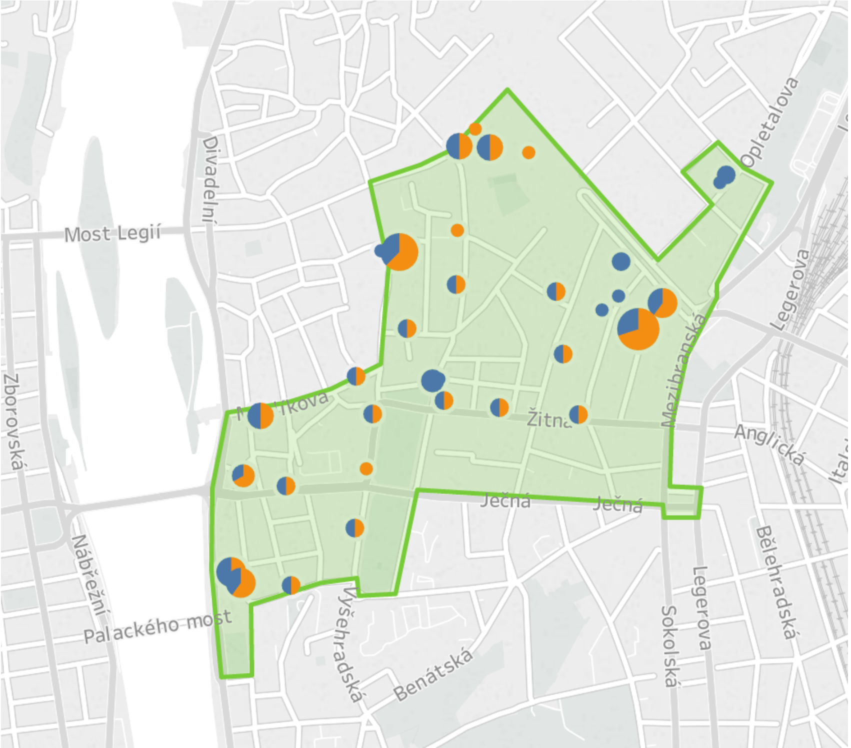


Figure 6.9: 0.8 km^2 large area covered by the geodatabase (with 51.9 km of pedestrian links), where the long-term experiment took part (green). Pies charts denote origin address (orange) and destination address (blue), their size denotes their frequency.

and receiving their payments.

6.6.4 Results and Discussion

During the 14 days of the experiment, we collected following data (see Figure 6.10).

Quantitative measures

Participants walked 44 routes with total length of 20.73 km altogether; 7 routes were unsuccessful (4 of these routes started or lead to pedestrian zones, which are hard to navigate in); 37 were successful, which results in a success rate of 84.01 % (92.50 % excluding 4 routes in pedestrian zones; what means only 3 unsuccessful routes). Participants walked 18 new routes in total (40.91 % of all routes).

A mean length of a route was 471.14 m ($SD = 263.01$), a mean duration of a route was 14.91 $minutes$ ($SD = 13.96$). A mean subjective judgement about the level of quality of both the navigation instructions and the routing for the routes was 1.81 ($SD = 1.17$, 95 % CI [1.45, 2.17]) on a 1-5 Likert scale (lower is better, see Figure 6.10 last row); a mean subjective judgement excluding all unsuccessful routes was 1.43 ($SD = 0.73$, 95 % CI [1.19, 1.67]). The participants traveled on average one route per day.

Daily diaries

Over the course of the experiment we collected 3 participants × 14 days = 42 daily diaries from the participants, 19 of which were filled-in. The participants did not use the navigation application prototype every day and thus the corresponding daily diaries were left empty. Many routes were to work, to a restaurant for a lunch or doing errands. However, the participants tried a lot of new routes out of curiosity. Some participants also tried an alternative route to their work using the navigation application prototype.

The participants' subjective judgment about the level of interaction comfort was 1.53 ($SD =$ 0.51, 95 % CI [1.28, 1.78]) on a 1-5 Likert scale (lower is better); subjective judgment about the level of assistance was 2.21 ($SD = 0.79$, 95 % CI [1.83, 2.59]) on a 1-5 Likert scale (lower is better, see Figure 6.10 first two rows).

Some participants reported reluctance to take out their phones when it was raining to control the navigation application prototype. On Android OS there was a problem with a screen reader and for one participant it was difficult to fill in the destination address. One participant also complained about the need to unlock the screen every time s/he wanted to interact with the application prototype. Altogether, the comments of the participants about the interaction comfort while walking were positive.

When the participants traveled through pedestrian zones, they often went astray (4 of 7 times). In the diaries, they mentioned that the navigation was most helpful for easier routes, however, it did not help them much in pedestrians zones. This was caused by insufficient data coverage in outlying parts of the experiment area. Some of the participants found it hard to answer the question about subjective judgment about the level of assistance on familiar routes, however, on some routes, they were certain it helped them.

5-day diaries

After the longer period of time, the participants mainly reported the inaccuracies in the geodatabase, which affected the quality of navigation instructions in the pedestrian zones or information about the length of pedestrian crossings. The most frequent critic was about the inability to change the route on the fly (e.g., automatic rerouting; in situations when the participants need to change the course of the route because of a construction zone, they need to fill in the origin address again), or automatic recognition of origin address (e.g., using GPS). Some participants reported that they found the experiment area rather limited. Some routes started from the same locations or ended in the same destinations as the routes walked previously, however, no route was walked twice by one participant, one route was walked once in both directions (see Figure 6.9).

The participants praised the feeling of independence (P2: *"When the map is without mistakes, it's comfortable and nice to walk a new route without preparation or asking passersby."*) when traveling alone or about details on a route as slope of sidewalk or traffic on pedestrian crossings (P2: *"The information about slope of the sidewalk and traffic information on pedestrian crossings are particularly useful."*). They also mentioned that the navigation application prototype helped them on longer routes, where they did not need to remember when to take the turn (P3: *"It's useful aid on routes I've never walked, I don't need to focus so much where*

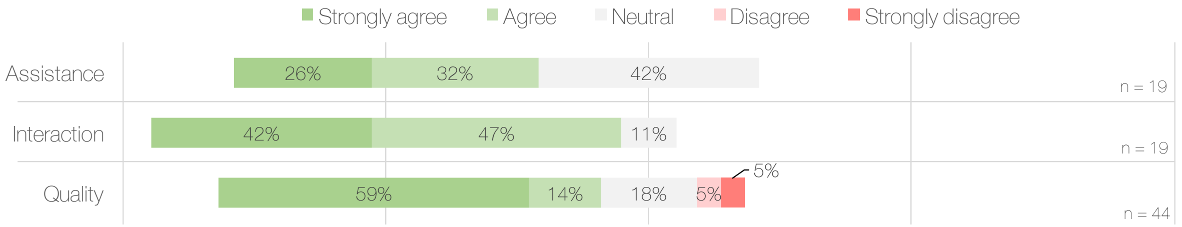


Figure 6.10: Subjective judgments about the level of assistance (summarized per day), easiness of interaction (summarized per day), and quality of navigation instructions (summarized over all routes), collected during the long-term diary study.

to turn and how many crossings I need to cross."). They would all recommend such navigation aid to their relatives or friends but with the warning about inaccuracies in the geodatabase.

Final report

In the final report, the participants expressed their concerns about collecting geographic data of a larger area of the city or even country into the geodatabase. One participant told us that s/he does not go new places as often and if s/he does s/he plans it in advance. Thus, s/he can use alternative services provided by local orientation and mobility specialist to get human-made route itinerary. Further, s/he was curious how would the system perform on more complicated parts of the city without the regular layout of blocks of buildings. Other participant missed GPS, which would let him/her know, whether s/he is proceeding correctly.

The long-term study showed that the navigation system with landmark-enhanced navigation instruction using the pedestrian link network PedestriNet is feasible even without usage of GPS (success rate 84 %). To improve the user experience of the accessible web navigation application prototype the feedback from the participants, such as simplified origin address input or usage of GPS for checking whether the traveler proceeds correctly, should be implemented. Finally, a larger area of the city needs to be covered by the pedestrian link network PedestriNet and additional improvements implemented into the geodatabase.

6.7 Essential Area Coverage

Production of geographical data is not trivial and is cost/time consuming task. To ensure cost-efficiency of deploying the system, we argue that only part of the city should be covered by the sidewalk network. We aim to identify of most interest among the local population of visually impaired people. To identify the area of interest of visually impaired pedestrians in the city center, we conducted a survey, to learn about their travel habits and travel related needs.

As the most popular means of transport among visually impaired people in the European context is public transport (47 % of trips [91]), we focused our research on usage of popular public transport stations in the city center.

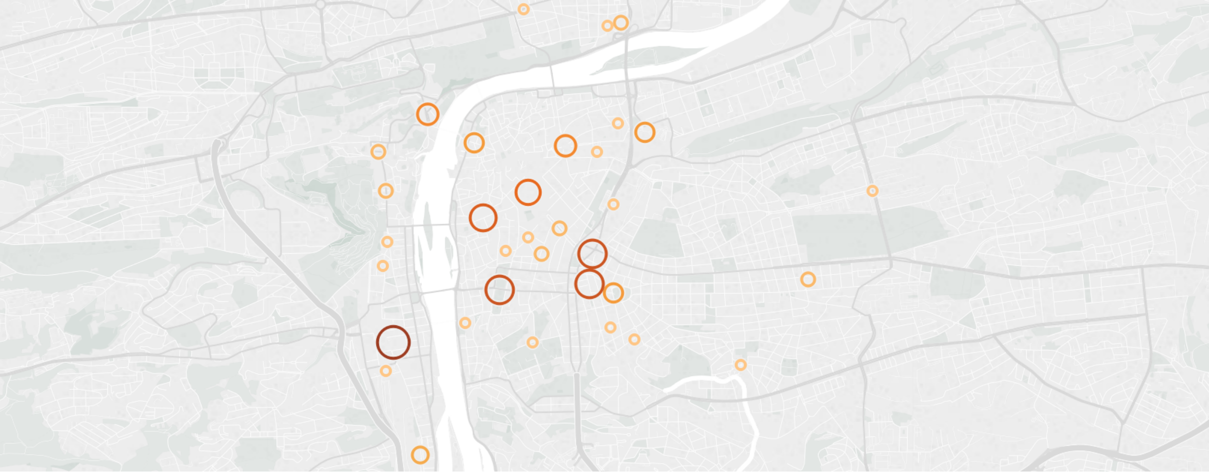


Figure 6.11: Data about the most used public transport stations (n = 18) by visually impaired respondents. Both size and color denote the frequency of the stations' usage. The map is cropped to show stations with the frequency of at least 2.

6.7.1 Participants

Eighteen visually impaired respondents (11 female, 7 male) were randomly recruited from the pool of long-term collaborators of our university. The respondents age ranged from 23 to 66 years ($mean = 36.78$, $SD = 12.30$). Six respondents had Category 4 visual impairment (light perception), 12 respondents had Category 5 visual impairment (no light perception) [113]. Fourteen respondents were congenitally blind, 4 respondents were late blind. All respondents were citizens of Prague, Czech Republic.

6.7.2 Apparatus

We used a standard smartphone with a headset attached to it to conduct the survey. The experimenter took notes for each question s/he asked directly into a spreadsheet table prepared in advance. The results were then analyzed and visualized using Tableau[5] visualization software.

6.7.3 Procedure

We conducted an over-the-phone survey with 18 respondents. The phone calls lasted from 5 to 10 minutes. The experimenter started with briefing and explanation of the survey purpose. Then the experimenter asked the respondents questions about navigation services they use and in which situations, their destinations in the city center, and public transport stations they use.

6.7.4 Results and Discussion

The following paragraphs summarize findings collected during the survey.

Public transport

The respondents mentioned fifty unique public transport stations where most of them were in the city center. One station was mentioned 12×, 3 stations 9×, 1 station 7×, 2 stations 5×, 3

---

[5] Tableau – www.tableau.com

stations 4×, 1 station 3×, 9 stations 2× and 29 stations 1×. We plotted the stations and their frequency onto a map (see Figure 6.11) to identify the area of the city, which is visited most frequently.

Types of destination

The most frequently used destinations reported by the respondents were shops (10×), work (7×), restaurant (7×), social gatherings (6×), culture (5×, doctor (4×), school (2×), library, church, tea-shop, and different kinds of offices (each 1×).

Using the data about types of destinations and frequently used public transport stations may help to identify the area, which is essential for efficient navigation of blind pedestrians in the city center. The collection of geographical data, the updating, the data distribution, and the financing depends on possibilities of particular cities that will be interested in systematic support of disabled citizens and visitors. The areas of primary importance are those where users would use routes connecting important points of interest (e.g., administrative offices, cultural objects, or transportation hubs).

6.8 Conclusions and Future Work

In this chapter, we report on the studies we conducted with 43 visually impaired participants.

We developed a method for automatic generation of landmark-enhanced navigation instruction for blind pedestrians, which is based on the creation of new sidewalk based GIS data structures, development of the algorithm for navigation instruction generation in natural language and routing algorithm optimized for needs of visually impaired travelers.

We conducted a comparative study of landmark-enhanced navigation instruction (*Landmark*) with metric-based navigation instruction (*Metric*) with 16 visually impaired participants. Although previous studies [88] show that landmark-based navigation is better for pedestrian navigation, the measured results were inconclusive. However subjective evaluation showed the preference of landmark-enhanced version. Moreover, we provide a set of design recommendations for the creation of navigation instructions related mainly to corners and crossings.

Furthermore, we investigated improved landmark-enhanced version (*Landmark*) in the qualitative study with 6 visually impaired participants resulting in additional recommendations related to leading lines or passages. The subjective evaluation indicated perception of *Landmark* method at high rate of safety, effectivity, and comfort.

To explore the usage of the system in practice, we conducted a long-term diary study with 3 visually impaired participants. The results showed promising success rate 84 % (even without using GPS) and overall satisfaction with the system, namely the improved feeling of independence mentioned by the participants.

Finally, we concentrated on identification of most frequently visited areas, which need to be covered by the geodatabase. We conducted a survey with 18 visually impaired respondents focused on usage of public transport stations, which seems like a promising method.

In the future, we will focus on integrating more landmarks such as recessed buildings or traffic sounds and opportunities for crowdsourcing of the geographical data (e.g., [121] or [86]). After integration of new landmarks, new evaluation has to be done to ensure the usability of the system. Further, the focus should be on providing users with possibility to change routing and to select from multiple routes by means of a dialog held during the navigation with the system.

6.9 Implications

In this chapter, we presented an iterative design towards landmark enhanced route itineraries inspired by the verbal description by contents and strategies used by visually impaired pedestrians while navigating over a phone (see Chapter 4). The evaluation of the efficiency and the qualitative aspects of the generated verbal route description show promising results. Having a tool for efficient natural language generation from GIS data, we may aim for systems which provide these verbal descriptions in various ways. There are more other options where to extend the work – namely to indoor, public transport and environment transitions (e.g., from indoor to outdoor) as in [11]. Furthermore, in this chapter, we provided answers to research question RQ5 (How will people with visual impairments perform with automatically generated verbal description during navigation?) as defined in Chapter 1. Detailed answers to the research questions are summarized in the chapter Conclusion 8.

In the next chapter, we utilize the results of Chapters 4 and 6 to model and prototype a conversational agent which using natural language interface provides the verbal description of routes in a city in a similar way the visually impaired person would do. Using natural language as an interaction technique enable us to design for visually impaired users with different navigation and orientation abilities and at the same time to provide personalized directions and level of detail of the verbal description.

Appendix: Random Variables

We classified the problems (or disruptions) observed during the Comparative Study (see Section 6.4) into four categories: Collisions with objects 27×, Interference of passerby people 11× (offering help 5× or grabbing and dragging 6×), Disruptions of the senses 17×, and Stress 3×. We consider Collisions with object and Interference of passerby people – offering help as common problems 31×. We consider Interference of passerby people – grabbing and dragging, Disruption of senses and Stress as serious problems 27×. Table 6.6 contains the full list of problems occurred during the comparative study classified into corresponding categories.

Category	Common problems	Serious problems
Collisions with objects	• stumbled over sun umbrella stand • 4× hit traffic sign • 3× stumbled over person • 2× stumbled over parking machine • 6× stumbled over dust bin • 6× stumbled over beer garden • 3× stumbled over car parked on sidewalk • stepped from sidewalk to road	• car on sidewalk (need to walk around on street)
Interference of passerby people	• 4× passerby offering help • passerby asking about the experiment	• 3× dragged by passerby • 2× passerby warning about obstacle • young passerby blocking the sidewalk
Disruption of senses	–	• 2× heavy rain (changed acoustic properties of environment) • road cleaning car (heavy noise) • ambulance passing by (heavy noise) • large group of tourists with suitcases (noise and crowd on sidewalk) • 2× damp roadway (louder noise of traffic) • 3× dustmen service (heavy noise) • car driving from passage (noise, danger) • 2× car with started engine on sidewalk (noise, danger) • reconstruction work on nearby building (heavy noise) • group of drunks (noise, danger) • 2× low sensitivity in fingers
Stress	–	• 2× nervous about the success of experiment • dropped white cane

Table 6.6: The occurrence of problems (disruptions) participants experienced classified according to type and severity.

Chapter 7

Conversational Agents for Physical World Navigation

Abstract. *This chapter presents a design process for developing a conversational navigation agent for visually impaired pedestrians where communication runs in natural language. This approach brings a lot of new problems with the opportunity to solve them in nontraditional ways. The conversation with the agent is an example of a problem-solving process with several complex parts of the solution needed to be executed by the user. The user can ask additional questions about each part of the solution, thus adaptively changing the level of detail of information acquired, or to alternate the whole process to fit user preferences. In this way, the agent can replace a human navigator. Using this design process exemplar, we provide guidance on creating similar conversational agents, which utilize a natural language user interface. The guidance is supported by the results of several experiments conducted with participants with visual impairments.*

Note. *This chapter is based on published book chapter Conversational Agents for Physical World Navigation [13].*

7.1 Introduction

The dominant user interface style today is WIMP GUI (graphical user interfaces based on windows, icons, menus and pointing devices) first developed by Xerox PARC in 1973 [7]. As the key interaction concept of WIMP GUI is direct manipulation, it requires fine motor manipulation with a pointing device and continuous visual control while manipulating a graphical object on the screen. Although the WIMP interface style was truly revolutionary at that time and allowed broad audience access to computers, it does not naturally support other forms of human interaction like speech, gestures, and facial expressions that systems are now able to interpret. In this way, WIMP GUI may limit interaction in new computing environments and various task domains, like 3D graphics, virtual reality environments, mobile interaction, and ubiquitous computing, where the primary task lies outside the computer system.

Since the 1990's we can observe a strong activity to define new interface styles, which will better fit this new situation. Dam [106] has introduced a Post-WIMP interface defined as an interface containing at least one interaction technique not dependent on classical 2D widgets such as menus and icons. Jakob Nielsen [74] came up with a term non-command user interfaces to characterize next-generation interfaces allowing the user to focus on the task rather than on the control of the user interface. His vision of the next-generation of UIs is syntax free, using a task-oriented approach instead of the traditional object-oriented one. This task-oriented approach specifies the unification of the object-action such that the user inputs a single token.

As the WIMP GUI puts stress on direct manipulation and visual control, it creates essential barriers for people with severe visual or motor impairment; this situation requires fundamental attention. But if we consider that using for example spoken dialogue, users with visual impairments can be even more efficient than the majority of the population by employing their highly-developed recognition abilities [118]. In the context of these concepts (Post-WIMP and non-command UI) an interface based on conversational agents [27] interacting in natural language is a promising approach for people with severe visual or motor impairment.

In this chapter, we introduce a design process used to develop a conversational navigation agent for visually impaired people. We will show how to design conversational agents for a complex knowledge-intensive problem solving process and how the problem solving process can be simplified using an on-demand level of detail approach.

The problem will be demonstrated on the following use-case:

> **Use-case:** Let us imagine a person with visual impairment wants to go from work to a new restaurant nearby. S/he uses conversational agent based navigation. The route is divided into separate segments – blocks of buildings, parts of a park, pedestrian crossing. At each of these segments, the person can ask the conversational agent questions about the length of the segment, leading lines, slopes, or the material of the sidewalk. However, on each segment s/he will (obviously) get a different answer. On the other hand, the person can at any time also ask a global question about the route such as how long is it, what is the estimated time of arrival, how many crossings are there, etc. Finally, the conversational agent may ask the person about his/her progress to get feedback that s/he is proceeding properly.

The activity of *problem solving* is defined by Lovett [65] as follows: “A problem occurs when there is something stopping you from getting from where you are at present to where you want to be – from your present state to your goal state – and you don’t know how to get around this obstacle.” This definition can be directly applied to the use-case. It is necessary to execute several nontrivial parts of the solution to find a way to the destination. There can occur various problems like determining location, finding the building corner, crossing the street, and choosing the right direction. In order to find solutions, a *problem space* – the space of all possible solutions to the problem – needs to be defined in advance [73].

In our use-case, the problem space is represented by a Geographical Information System (GIS), a knowledge base containing all possible routes (see Figure 7.5). The route in the use-case is a result of problem solving process, a solution to a problem of finding a route from A to B based on user preferences. Each segment of the route is representing a part of the solution.

There is an interesting conceptual parallel between problem space (represented by the GIS) containing all possible routes the user can walk and *conversation space* containing all possible conversation paths about the route. In our case, the conversation space is automatically generated from the GIS based on the solution found in the problem space.

Following on the above use-case given, at each segment of the route, the user gets a general description of the task, which has to be executed to proceed to the next segment. Then the user can naturally ask about various attributes of the current segment, which will help to execute the task. The user asks the questions with various intents – in our use-case, to get information about distance or position of leading line – and the conversational agent has to be prepared to provide answers, which differ (different length, leading line) in each segment. The important role of the HCI designer is to determine the user intents in advance in order to let the conversational agent react to a majority of relevant ones and to provide fallback scenarios to respond to unknown intents.

Similarly, the conversational agent itself can ask the user about his/her progress in executing a part of the solution by the conversational agent itself. In this way, the user collaborates with the conversational agent introducing a mixed-initiative interaction. In our use-case, the agent asks the user about the progress when executing particular segment by means of verifying various properties of the environment. As the same properties have different values for different segments, the agent needs to determine the user location and direction of his/her movement. One way is for the agent to ask a closed-ended question (yes/no), which abstracts the user responses from particular context-dependent properties (i.e., a position of a landmark). On the other hand, when the agent asks an open-ended question, it would provide the user with more freedom and comfort. However, open questions bring complexity to the course of the dialogue – HCI designers need to account for many possible context-dependent user responses, which in some case would be impossible to abstract – to enable the agent to react to whichever value of particular property user answers.

Let us imagine conversational a system asking about the slope of the sidewalk. The first yes/no approach would result in the question "Is the sidewalk downhill?" with only two possible responses "Yes" or "No". The agent can extract a value (downhill) for the slope property from the knowledge base in advance and only has to recognize the yes/no answer. This approach avoids more specific or out-of-scope answers, which can lead to uncertainty while choosing the answer ("Yes, but..."). On the contrary, the second approach would result in the question "What is the slope of the sidewalk?" which has a wide range of possible user responses, like "Downhill" or "It's slightly downhill" or "Approximately 10 degrees", etc. In this case, the user can answer more freely with a higher confidence of providing accurate information. However, the conversational agent cannot map all the possible answers to values of the particular property stored in the knowledge base (see Subsection 7.3.2).

The knowledge base (GIS in navigation use-case) is an inseparable part of the conversational agent also in terms of personalization. In a use-case of navigation, different strategies for finding a solution to a navigation problem in the GIS can provide a personalized route for users [108]. Furthermore, the knowledge base provides the user with a context. In our use-case for example, providing detail information about the crossroad helps to plan possible alternative routes. At this point the user may be unsatisfied with the solution provided by the conversational agent. Knowledge of the crossroad opens a possibility to change the route if the environment or user preferences change. These places (e.g., crossroad) are typically called *decision points*. Taking decision points into account when designing conversational agent provides the user with more freedom and lets them decide which way of proceeding is best for them (see Subsection 7.3.3).

For an HCI designer of a conversational agent trying to design this kind of interaction, a set of important questions arise: How to integrate dialog with knowledge base? How to define user intents applicable to all parts of the solution? How to control the progress of user on different tasks or how to manage changes of solution (human error, preference, habits)? The following sections aim to provide answers to these questions. The design process will be illustrated on an example of designing a conversational navigation agent for visually impaired users.

7.2 Related Work

This section provides the background to the presented use-case in the context of conversational user interfaces design.

Conversational user interfaces are in general useful in situations when the user has limitations of some of the communication/interaction modalities, typically sight or touch. For example, a driver is limited in sight when interacting with some kind of device as s/he has to pay attention to the roadway, similarly a visually impaired person cannot use the display of the device and has to use the hearing modality. On the other hand, sometimes, it is not possible to touch the device to interact with it – for example when a person is cooking and has dirty hands, similarly a person with limited upper hand dexterity (quadriplegic) cannot use touch and has to use a different modality, in this case, voice.

The *conversational agent* is a dialogue system composed of natural language understanding module (understanding user intents), dialogue management module (defining the course of the dialogue), natural language generation module (generating meaningful instruction and information needed to proceed to next part of the solution), providing a natural language user interface to the user (see Figure 7.1). All of these components are further subdivided according to needs of developers, designers, researchers and intended users [27]. HCI designers and researchers can contribute components, which are responsible for shaping the course of the dialog (dialog management) and the verbal output (natural language generation, see Figure 7.1).

7.2.1 Inspiration for Design

For a better understanding of the use case, we provide insights into related navigation aids used and difficulties encountered by visually impaired people. The ability to travel independently to desired destinations is required for a satisfactory level of quality of life and self-confidence, however, people with a severe visual impairment struggle with their mobility and travel-related activities are significantly reduced [48]. Although visually impaired people undergo special training to learn specific navigation and orientation skills, techniques and strategies, 30 % of them never leave their home alone without a sighted guide [28], [112].

The inspiration for the conversational approach in the navigation of pedestrians with visual impairments surfaced from an existing solution: tele-assistance centers for the navigation of visually impaired users. Tele-assistance centers are operated by professional orientation and mobility specialists, who identify the position of the visually impaired pedestrians (either by a camera attached to the user [23] or purely by verbal description [109]), can provide suitable and efficient verbal navigation. The tele-assistance centers have two scalability problems. First

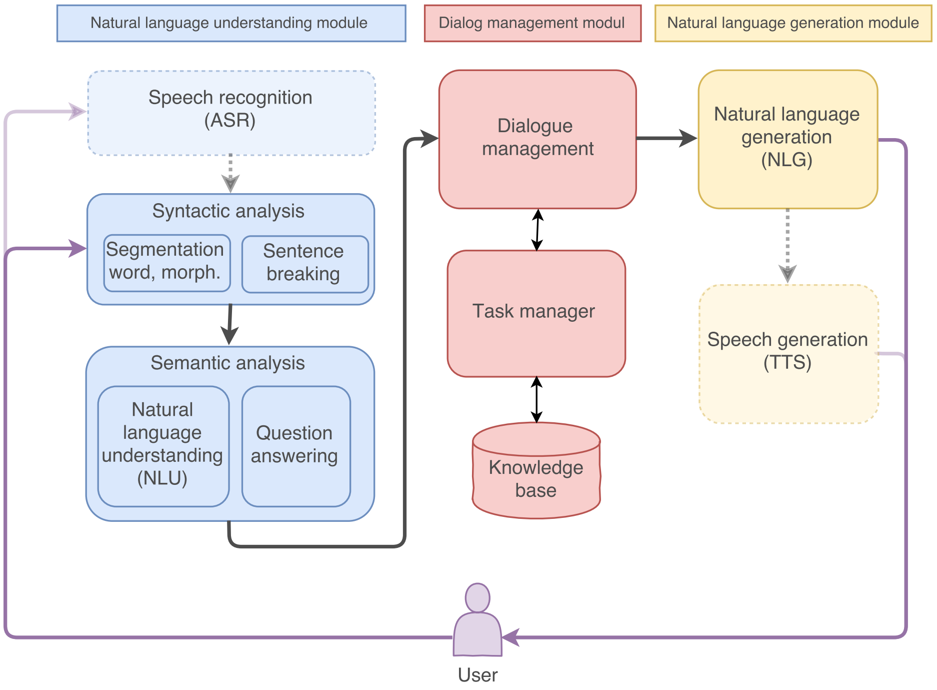


Figure 7.1: The conversational agent typically consists of a natural language understanding module, dialog management module and natural language generation module [59]. The connection to external knowledge systems (like GIS) is realized by the Task manager.

is the limited operation area, as the gathering of a suitable set of landmarks used in verbal descriptions for a particular area often requires the operator's physical presence on the spot. Second is the limited operating hours due to use of human operators.

Verbal navigation also appears in books about typhlopedy[1] such as Welsh [114], who recommend that verbal description of a route be enhanced with landmarks (e.g., corner, crossroad) in a form of pre-prepared route description. Similarly, experiments of Bradley and Dunlop [20] showed the preference of verbal descriptions of a route recorder by visually impaired users over those recorded by sighted ones.

First attempts of dialog based navigation of the visually impaired dates back to 1996 with Strothotte et al. [97] whose dialog system enabled visually impaired users to ask basic questions about estimated time of arrival, orientation cues or approximate location.

Use of a conversational agent is also related to engagement and creation of cognitive maps. When people use an automatic global satellite navigation network based application, they are virtually dragged through the calculated route without having a chance to activate their wayfinding and spatial cognitive functions [61]. This leads to degradation in spatial knowledge acquisition [76]. When interacting more with the environment and landmarks along the route, the user is more engaged and creates a better understanding of the surrounding environment.

7.2.2 Natural Language Generation and Interaction

To enable the conversational agent to act in a pleasant and believable way from the user's point of view, attention needs to be paid also to natural language generation. When focusing on

---

[1] Typhlopedy deals with special pedagogical care for people with visual impairments.

problem solving and natural language generation we look at natural language generation from open domain knowledge. Examples of this are generating natural language from Wikipedia, and letting the user navigate within information and change topics [115].

If we look back at the use-case of navigation of visually impaired users, what is important in the context of navigation are details such as landmarks, which proved to be very frequent and useful [63], [67], [88]. However, including many details in the generated text means that the person has to listen longer to the screen reader and needs to remember far more information [12]. So our approach is in line with "preference for minimisation" observed in natural human conversations [89], where speakers design their talk to be the most efficient. In other words, dialogue can allow more complex information to be conveyed compared to traditional user interfaces using a single utterance [2].

As described by Allen et al. [2]: "Dialogue-based interfaces allow the possibility of extended mixed-initiative interaction [3], [26]. This approach models the human-machine interaction after human collaborative problem solving. Rather than viewing the interaction as a series of commands, the interaction involves defining and discussing tasks, exploring ways to execute the task, and collaborating to get it done. Most importantly, all interactions are contextually interpreted with respect to the interactions performed to this point, allowing the system to anticipate the user needs and provide responses that best further the user goals. Such systems will create a new paradigm for human-computer interaction."

In other words, the mixed-initiative interaction provides a natural way to let conversational agent collaborate with the user rather than passively accepting simple commands.

7.2.3 How to Start

When designing a conversational agent, initial research when designing a conversational agent usually starts with experiments involving two humans having a dialog about a problem which involves completing a nontrivial task, typically in a form of a Wizard of Oz experiment [41]. Examples of such research can be found also in Vystricil et al. [109] and Balata et al. [9] where people with visual impairments navigated each other in pairs in unknown environments; one of them (*navigator*) was familiar with the environment, whereas the other one (*traveler*) was new to the environment.

The initial research used to design for our use-case – conversational navigation for visually impaired pedestrians – also started with Wizard of Oz experiments (see Section 7.3)

7.3 Designing Conversation for Physical World Navigation

This section provides details of designing conversation in knowledge rich contexts. The design process will be demonstrated on the conversational agent "NavAgent" for physical world navigation of users with visual impairments.

Excerpt 7.1 illustrates an interaction with current navigation systems (e.g., Google Maps). The initiative is only on the system's side (A), and the user (U) is passively accepting to commands.

On the other hand, an example of the mixed-initiative conversation with NavAgent (A) from our use-case (see Section 7.1) can be found in Excerpt 7.2. The main difference between the

01 A: Turn to the left and go to the corner.
02 A: Continue straight and cross the street.
03 A: Turn to the left and go to address five
04 Main Street.

Excerpt 7.1: System Initiative

two approaches is highlighted in Excerpt 7.2 in bold. Namely at lines 2-7 and 9-14 where the user (U) asks additional questions which would not be possible to ask in current voice-based navigation systems.

01 A: **Turn to the left and go to the corner.**
02 U: How far is it?
03 A: It is approximately 150 meters.
04 U: What is the shape of that corner?
05 A: There will be a round corner.
06 U: Okay, I am here.
07 A: Okay, I'll tell you how to continue.
08 **Continue straight and cross the street.**
09 U: Tell me about the traffic.
10 A: It is a crossing with tactile and sound
11 signalization, with one-way traffic
12 from the right.
13 U: I'm on the other side.
14 A: Okay, I'll tell you how to continue.
15 **Turn to the left and go to address five**
16 **Main street.**

Excerpt 7.2: Mixed Initiative

At line 1 in Excerpt 7.2, the agent provides basic information about the segment user needs to execute. At lines 2-5 the user asks about the properties of the segment – in this way the user selects the needed level of detail to complete the segment (part of the solution) – such as distance or shape of the corner. By using this on-demand level of detail technique, it is possible to avoid overwhelming the user with unnecessary information and save time and effort needed to listen to and remember long and detailed navigation instruction. Then, the initiative is fully on the user side. At line 6 the user confirms completion of one segment and then information about the next segment is presented. At lines 9-12, the user asks again about the additional properties (i.e., landmarks) of the segment. At line 13, the user confirms the segment completion again. The interaction continues in Excerpt 7.3.

At line 17 in Excerpt 7.3, the initiative switches to the agent. Here, NavAgent asks the user about the progress of the navigation step using a property (a leading line – building) with attribute (position – right-hand side) from the knowledge base using a closed-ended question. First, the user does not know (line 19) and is prompted again with guidance how to get the information safely (lines 20-22). At line 23, the user confirms a position of the leading line. Line 24 shows grounding of the agent.

At lines 25-27, the initiative is taken again by the user, who asks about points of interest around the block of buildings to get better contextual information for forming a cognitive map.

```
17  A:  Do you have the buildings on your right-
18      hand side?
19  U:  I don't know.
20  A:  Please stop walking and try to explore your
21      surroundings.  Do you have the buildings on
22      your right-hand side?
23  U:  Yes.
24  A:  Great.  You can continue.
        ((lines omitted))
25  U:  Is there any restaurant?
26  A:  There are a grocery store, a travel agency
27      and Big Belly Burger restaurant.
```

Excerpt 7.3: Location Check

7.3.1 User Intents

As shown in Excerpts 7.2 and 7.3 above, the user can ask questions about various properties of the segment. To define the set of user intents and the structure of the dialog, we conducted a user study with people with visual impairment [9]. We explored a situation when the blind *navigator* (*navigator*) forms a natural source of suitable landmarks with their descriptions and with routing strategies optimized for blind travelers (*traveler*). During the experiment, in 19 sessions, we observed the *navigator* guiding the *traveler* via phone call in an urban environment (see Figure 7.2), we gathered a set of problems that occurred during the navigation, and we classified the problems into activities performed by the *navigator* and by the *traveler*. In the end, we identified the activities of the *navigator* and the *traveler*, their turn-taking, and grounding strategies. Finally, we transcribed and categorized all dialogs, and identified queries made by the actors. The dialog between the blind *navigator* (N) and the blind *traveler* (T) looked like the one in Excerpt 7.4:

```
01  N:  Well, you'll go approximately fifty meters.
02      There are some shop doors open along the
03      sidewalk, and after the fifty meters, there
04      is different paving - cobblestones.
05  T:  Will I go downhill again?
06  N:  Yes, you'll go downhill and after that fifty
07      meters, there will be also a large door at
08      that place with, cobblestones
```

Excerpt 7.4: Slope Inquiry

At lines 1-4 in Excerpt 7.4 the navigator provides information about the segment that needed to be executed – namely distance, points of interest around the building block and landmark represented by the change of sidewalk material at the segment. Then the initiative changed and traveler (at line 5) asks about the slope of the sidewalk. At lines 6-8, first, there is a grounding and providing of properties of the sidewalk – namely the slope of the sidewalk and again landmark at the end of the segment. In Excerpt 7.5 the dialog continues.

At lines 1-2 in Excerpt 7.5, the *traveler* confirms the landmark at the end of the segment – big door. Then, the *navigator* provides information about the next segment (lines 3-5). At lines 5-6, *traveler* performs grounding and asks for confirmation. At the last lines, the navigation provides grounding and adds additional properties to the segment.

```
01 T: Okay, there are some dustbins and a big door
02    Great.
03 N: Well, turn so that you have the door at your
04    back and then cross the street.
05 T: Turn my back to the door and cross the street,
06    yes?
07 N: Yes, and keep slightly to the left, there are
08    sometimes cars parked.
```

Excerpt 7.5: Understanding Check

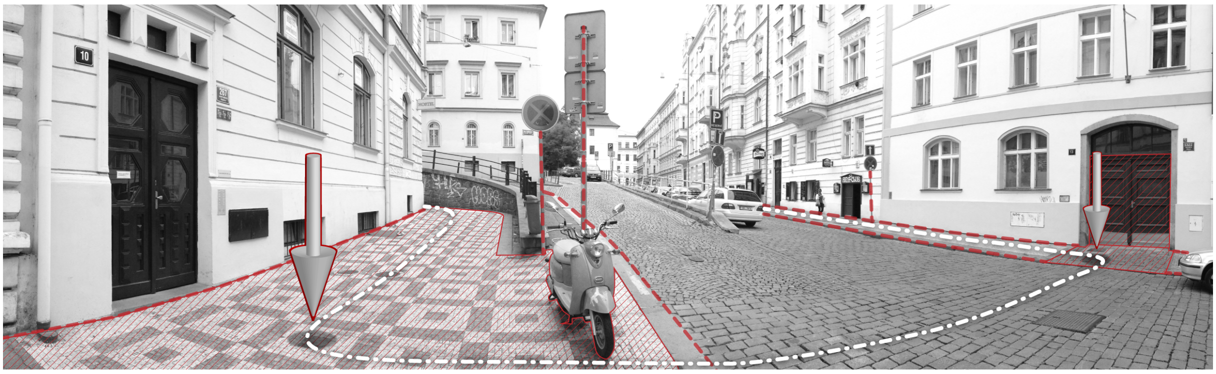


Figure 7.2: The urban environment where the experiment with *navigator* and *traveler* took place. Thick red dashed lines identify landmarks (position of building, curb, traffic signs poles); red hatched areas denote area landmarks (on the right different paving and doors, on the left slope of the sidewalk); large white arrows denote decision points; while dot-and-dash line denotes the path walked by *traveler* from right to left.

In general, we identified the following classes of activities: “Navigator describing the environment”, “Navigator giving navigation instructions”, “Navigator determining *traveler’s* position”, “Traveler executing navigation instruction”, and “Traveler identifying landmarks”. These activities were next used as cornerstones for designing intents users may have when interacting with the conversational agent. For the design of the user intents, we used only activities related to the *traveler* in the Excerpt 7.4 and 7.5:

- *Traveler* executing navigation instruction (i.e., lines 5-6 , Excerpt 7.5)
  - intents: distance request, direction check, completion report, previous request, help request
- *Traveler* identifying landmarks (i.e., line 5, Excerpt 7.4 and lines 1-2, Excerpt 7.5)
  - intents: end landmark information request, shape information request, material information request, traffic information request, leading line information request

The intents designed based on observed activities from class “*Traveler* executing navigation instruction” identified in Excerpt 7.4 and 7.5 correspond to lines 2, 6 and 13 in Excerpts 7.2. Intents designed based on observed activities from class “*Traveler* identifying landmarks” identified in Excerpt 7.4 and 7.5 correspond to lines 4, 9 and 25 in Excerpt 7.2 and 7.3.

The selection of the intents has to be done with respect to the knowledge base structure (GIS in our case). The GIS contains only some information about the environment (i.e., it omits information about dustbins mentioned in Excerpt 7.4 by the *traveler* or big door mentioned by *navigator*) like shapes of corners, sidewalk material, traffic information, etc. If the user mentions an unknown landmark "unknown landmark report" intent is used. Further, the general intents have to be added, such as "previous request, "disagreement expression", "agreement expression" (both used during grounding), "repeat request", "small talk expression" (to ask user to get back to navigation, see subsection User Experience) and "anything else" (to provide response to unrecognized intents).

Many natural language understanding modules need sample data for each intent to be recognized. Here, usually, a researcher or a domain expert comes in place and provides basic sample data based on previous research or experience. However, when data from real human-human interactions are available, they should be included too. In this case, data from Excerpt 7.4 were used to improve sample data (more in Section 4); transcribed utterances of *travelers* representing particular intents were added as samples to natural language understanding module.

7.3.2 Dialog Structure

Continuing on with Excerpts 7.4 and 7.5, we used the following activities of the navigator to design intents of the conversational agent and mixed-initiative interaction.

- *Navigator* describing the environment & *Navigator* giving navigation instructions (i.e., lines 1-4, Excerpt 7.4)
  - intents: providing environment description with action to perform to complete the segment
- *Navigator* determining *traveler's* position (i.e., line 5, Excerpt 7.4 and lines 1-2, Excerpt 7.5)
  - intents: general progress request, landmark properties request

The intents designed based on observed activities from class "*Navigator* describing the environment" & "*Navigator* giving navigation instructions" identified in Excerpt 7.4 correspond to lines 1, 7-8, 14-16 and 24 in Excerpt 7.2 and 7.3. Intents designed based on observed activities from class "*Navigator* determining traveler's position" identified in Excerpt 7.4 and 7.5 correspond to lines 6, 13 and 17-24 in Excerpt 7.2 and 7.3.

As you can see at lines 17-24 in Excerpt 7.3, the conversational agent tries to get the answer from the *traveler* even if s/he does not know. Similarly, there is a correction mechanism when the users proceed incorrectly (i.e., the building should be on the right, but the user says it is on the left).

From the same study as Excerpt 7.4 and 7.5, we observed the structure of the dialog. We mainly focused on the successful strategies of the *navigator* and *traveler*, which resulted in successful completion of the route (correct execution of all segments). Namely, these were those where both *navigator* and *traveler* paid more attention to the properties of the environment (i.e., landmarks, see Figure 7.2) than those where they dominantly focused on directions and distances. When designing the dialog model, we tried to simulate successful strategies (i.e., activities *Navigator* determining *traveler's* class – questions like "How's it going?", or "Is everything

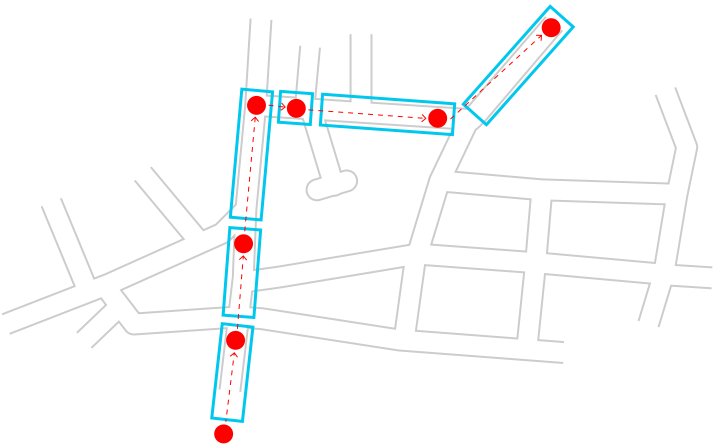


Figure 7.3: Route segments in navigation from A to B. A route represents a solution to a problem, where segments are parts of that solution. By executing the segments, the user solves the problem.

O.K.?" were useful and frequent, similarly more concrete ones like "Do you have buildings on your left-hand side?") as well as strategies employed when the *traveler* went astray (i.e., backtracking to last successfully completed segment).

As the context is different for each segment of a route, we separate it from the dialog system. The dialog model thus works on an abstract representation of the knowledge base properties. Instead of hardcoding "It is approximately 150 meters." in the dialog model, there is a "DISTANCE" keyword, replaced outside of the dialog model with appropriate information based on the information from the knowledge base. We used this strategy correspondingly to all other users' intents. In this way, we designed a dialog model, which is abstract from the concrete segment and can be generalized to all possible segments (see Figure 7.3) as part for all solutions of the problem solving process (see Figure 7.4).

Usage of abstraction in a dialog model allows us to change the solution in any step of the problem solving process. In this way, it is possible to react to context-related changes from the user preferences (e.g., when it begins raining user prefers usage of public transport instead of walking). For more details see Section 7.3.3.

Most of the time, the initiative is on the user side, and the conversational agent acts more as a question answering system [49]. The conversational agent takes the initiative after ⅓ and ⅔ of the approximate time needed to complete a segment except those which are too short (such as pedestrian crossings), as seen at lines 17-24 in Excerpt 7.3. Questions about general progress and landmark properties are shuffled to let the conversational agent feel more natural.

7.3.3 Alternative Routes

When the problem-solving process is running in a knowledge rich context, the user preferences are tightly related to it and become context-related. These preferences are given for the resources the user has at his/her disposal (e.g., availability of public transport), or by user skills or expertise. Allowing the user to take alternative routes, i.e., multiple trajectories in the problem space, let the conversational agent feel more natural and its interaction less rigid.

In the use-case (see Section 1) the context is the route, and its surrounding environment and

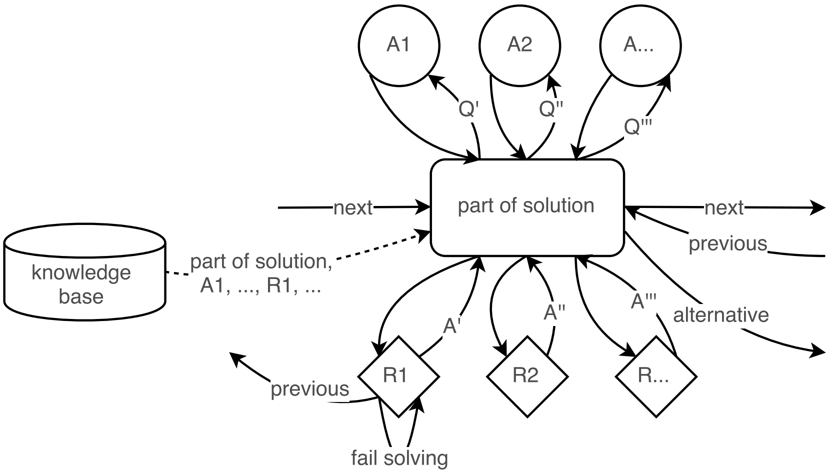


Figure 7.4: Generalized dialogue model for each part of the solution – Q', Q'', Q''', ..., Q^n are user's questions, A1, A2, ..., An are agent's responses, R1, R2, ..., Rn are agent's request on the user, A', A'', A''', ..., A^n are user's response to the system. Agent's responses A1, ..., An and agent's requests R1, ..., Rn are generated from the knowledge based independently of dialog model.

the user preferences can be roughly grouped into following categories:

- *route preferences* – length, number and type of crossings, type of traffic, slope and surface of the sidewalks, weather, etc.;
- *navigation instruction level of detail preferences* – familiarity of the environment;
- *conversation style preferences* – level of formality, frequency of grounding, initiative of the system.

The design issue is how to determine these context-related preferences. A common method, asking the user to fill out a form with many questions, will not work properly. The preferences cannot be described in a general way (e.g., "user prefers crossings with signalization"; "user prefers the shortest route") as they are influenced by the whole context, i.e., preferences in current situation are influenced by past and upcoming situations (e.g., previous route segment and next route segment influences preferences related to current route segment in which the user is located). For example in our use-case, avoiding several small crossings without signalization at the cost of necessity to cross the straight street several times to meet the preference "crossing with signalization only" is obviously not a good solution.

There are crudely two different approaches that can be used for obtaining the user preferences and their relation to the context (e.g., route and its surrounding):

1. individual observation of user behavior in various situations;
2. implementing a specific conversation with the user related to the preferences, habits, and values.

The first approach is easy to design, however, it is time consuming and costly as it requires an excessive amount of manpower to observe and annotate users' behavior. The second approach is hard to design as we are dealing here with a conversation trying to determine the user values and generalized needs which may lead to very complicated dialogue structures, which becomes quickly overwhelming and confusing. Moreover, apart from the objective preferences like distances or types of crossings, users' preferences may be connected to particular places in the environment or users' familiarity with the environment.

Our method divides the process into two phases (it is designed in a collaboration with TU Delft and based on Pasotti et al. [77]). The first time the user uses the conversational agent, s/he goes through a learning phase, where the agent tries to create a knowledge structure about his/her values and habits (e.g., preference towards taking safer crossings when traveling in unknown places or preferences towards walking more frequently to work than taking a bus). The conversational agent asks questions exploring ways of doing and parts of doing some travel-related activity (see Excerpts 7.6 and 7.7). Further, it asks a question about frequencies, values and habits related to the activities (see Excerpt 7.8). The knowledge structure is later used to find a solution to a navigation problem. Excerpts 7.6 and 7.8 are transcribed from the initial user study with six participants with visual impairments.

```
01  A:  Name several daily or weekly activities related
02      to independent traveling in a city with which
03      I can provide you an assistance.
04      (7.0)
05  U:  Well, shopping, traveling by subway, getting
06      information about public transport lines.
07      How much should I say?
08  A:  I’s enough.
09  U:  Okay.
10      ((lines omitted))
```

Excerpt 7.6: Inquiry of initial list of activities

At lines 1-3 in Excerpt 7.6 the conversational agent asks introduction question and inquiries list of travel related activities, here the user hesitates (the silence) before providing the list (lines 5-7). At that time s/he is not sure whether how much s/he should mention at the beginning.

```
11  A:  What do you do as parts of shopping?
12  U:  Searching of the counter, dictating the grocery
13      list, waiting in the line, searching the exit.
14  A:  Is searching of the counter part of shopping?
15  U:  Yes.
16  A:  Is dictating the grocery list part of shopping?
17  U:  Yes.
18      ((lines omitted))
```

Excerpt 7.7: Parts of inquiry

At line 11 in Excerpt 7.7 the agent requires a list of parts of the activity mentioned at the beginning, followed by explicit confirmation of the items at lines 14-18 (explicit confirmation concerned some users, however, some of the users were happy with it).

Further, the conversational agent asks the user about the values certain activity promotes or demotes. At line 19 in Excerpt 7.8 the agent asks the question, however, the user hesitated for a period of time so the agent tried an alternative question (line 22). The user continues with an exhaustive list of values (lines 23-26).

```
19  A:  What values do you feel shopping promotes
20      or demotes?
21      (10.0)
22  A:  What important life goal does shopping support?
23  U:  Hm, well, it satisfies my needs, full-fills my
24      needs for goods, it denotes patience, it
25      denotes social contacts.  It promotes my
26      orientation skills and stress resistance.
```

Excerpt 7.8: Values inquiry

Eventually, the user has the opportunity to change the solution (i.e., route) based on his/her preferences or habits at each decision point along the route – Let us imagine a user, who always wants to avoid parks. However, s/he likes one particular park in which it is easy for him/her to navigate. When s/he is guided along the park, s/he can tell the agent at nearest decision point that s/he want to change the route and go (for example) left instead of right. The agent will try to understand the reasons behind this decision, and it will add the local preference to the knowledge structure accordingly.

On the other hand, the agent can not only learn habitual preferences for particular places in the environment, but it can also support the user towards more appropriate behavior – Let us imagine a user, who is guided in a city. When the agent asks him/her to turn left, s/he can tell the agent s/he disagrees, and that s/he wants to use a different route and continue straight. In this case, the agent will try intervene, explain that the habitual route leads over dangerous crossing and encourage the user to take a safer route.

Similarly, for a sudden change in the context, another user prefers walking instead of using public transport. However, when it is raining, s/he want to use public transport, and the route has to change accordingly, and the agent has to notify the user about the change of condition so that s/he will understand, why the route is different from usual. Then, the user can accept or reject the newly proposed route.

Taking alternative routes and using users' habits and values to customize the preferences is still an unsolved problem and currently, we are dealing with early stages of Wizard of Oz evaluation of dialog models for the first phase (building knowledge structure).

7.3.4 Natural Language Description

In this subsubsection we provide description on how we created the module used for natural language generation of descriptions of segments as can be seen in Excerpt 7.2, at lines 1-5.

To generate natural language from the knowledge base using GIS, new algorithms had to be developed to create a description of any possible route in a way that mirrors those prepared by orientation and mobility specialist for the visually impaired.

By using specific data structures (see Figure 7.5) and algorithms we addressed the issue of time requirements (the route description generated immediately) and rigidity (user can select whichever origin and destination s/he wants) and evaluated the efficiency of such route description in comparison to commercially available navigation systems in a series of experiments [12] (see Chapter 6). The generated segment description can be found in Excerpt 7.9.

The description of a route as in the example above (Excerpt 7.9) contains all the information

```
You are at the beveled corner of Main
Square and Central Avenue.
Turn to the left and go approximately
one hundred meters slightly downhill to
the corner with Oak Street.
The street bends to the left.  Keep the
buildings on your right-hand side.  The
sidewalk is made from small paving block.
```

Excerpt 7.9: Full Segment Description

needed for successful completion of a particular segment. However, there is a problem of length of the text – time to listen and amount of the information needed to be remembered. In the conversational agent (Excerpt 7.2 and 7.3), we use the same approach used to generate the whole description of a segment at once. We created templates of utterances for all possible situation which exist in the problem space (GIS). Some of them are a concretization of another one, when there is more information about a particular segment or when based on the context some are more important than the others (i.e., to mention a corner of a block of buildings proved to be more efficient than to mention crossing, which is on the same spot).

When a route is found by the algorithm in the GIS, all responses for user intents are generated in advance for each segment and saved into an array (see Excerpt 7.10). Then according to user's intents, the agent provides the appropriate answer.

```
   SLOPES =  [``It is slightly downhill.'',``It is
             flat.'', ``It is slightly uphill.''];
DISTANCES =  [``It is approximately 150 meters.'',
             ``The crossing is approximately 5
             meters long.'', ``It is approximately
             40 meters.'']
```

Excerpt 7.10: Agent's utterances

To provide the agent with the right answer for the user, there is a task manager (see Figure 7.1), i.e., a middleware JavaScript application, which handles connection to the knowledge base, records information about current segment and connect to the dialog management server.

7.3.5 User Experience

The HCI designers of the conversational agent have the possibility to shape its virtual personality to their particular needs. The dominance of the agent can be used to have the system which is more persuasive, or, on the other hand, one may need a funny agent which entertains the users.

From the Excerpt 7.2 and 7.3, we can identify the following UX design approaches. The entity called NavAgent uses informal addressing (e.g., German "Du" French "Tu") to communicate with the user. It should induce a feeling of briskness of the navigation instructions and makes the whole application more personal (i.e., "preference for minimisation, see Subsection 7.2.2). We also trained the natural language understanding module to recognize a small talk. However, when the user starts small talk, NavAgent responds sharply to get back to the navigation task, which should emphasize the serious matter of navigation.

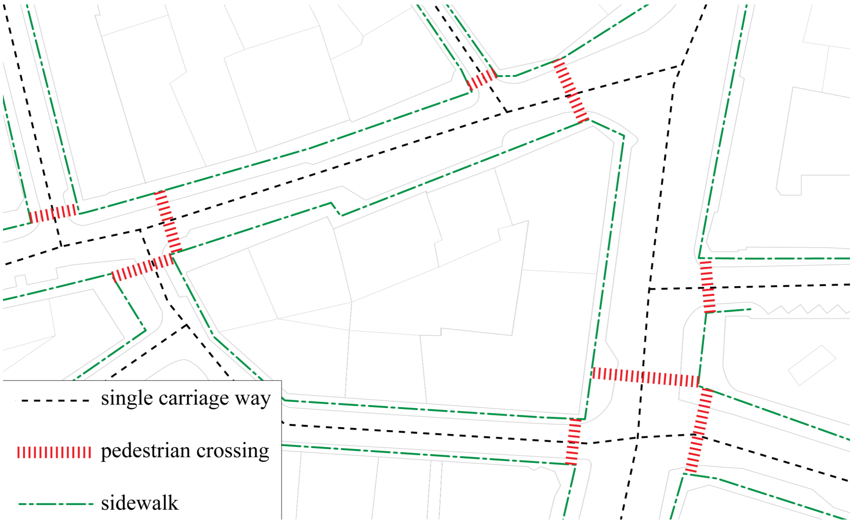


Figure 7.5: GIS used to generate natural language description of navigation steps provided by the conversational agent. Black dashed lines represent street network, green dot-and-dashed pavement network, red thick dashed line pedestrian crossings.

Time to time, NavAgent proactively asks users about their progress (see Subsection 7.3.2), which intends to both get their confirmation of location and to make them feel that NavAgent cares about them and their safety and well-being. The user can ask for repetition of any NavAgent's utterance based on last recognized intent.

Even though, the non-verbal sounds and interaction using natural language form an essential modality enabling visually impaired and blind people to interact with the world around them, users can interact with the conversational agent either by dictation or typing their queries and responses depending on their preference and context.

7.4 Iterating on Conversational Design

7.4.1 Evaluation design

The first steps we took to explore the design problem space for the conversational agent were based on observation of human-human conversation. In the first experiment, a human expert navigated participants with visual impairments. The purpose was mainly to explore the feasibility of conversational navigation and to observe language used by participants for different contexts (crossing street, walking in park or street or indoor) [109].

Further, an experiment involving two participants [9] from the user group of people with visual impairments was conducted to observe language and strategies of both actors of the conversation. From the second experiment, we gained useful knowledge about the dialog structure and intents together with training data.

The analysis started with the data collected from the second experiment. First, we analyzed the navigation strategies participants used to see which ones led to success and which ones led to failure. We placed the findings into the context on a hand-drawn map of the environment where the experiment occurred (see Figure 7.6). The analysis helped us to sketch the first version of dialog structure. Then, we transcribed the conversation, identified the intents and entities

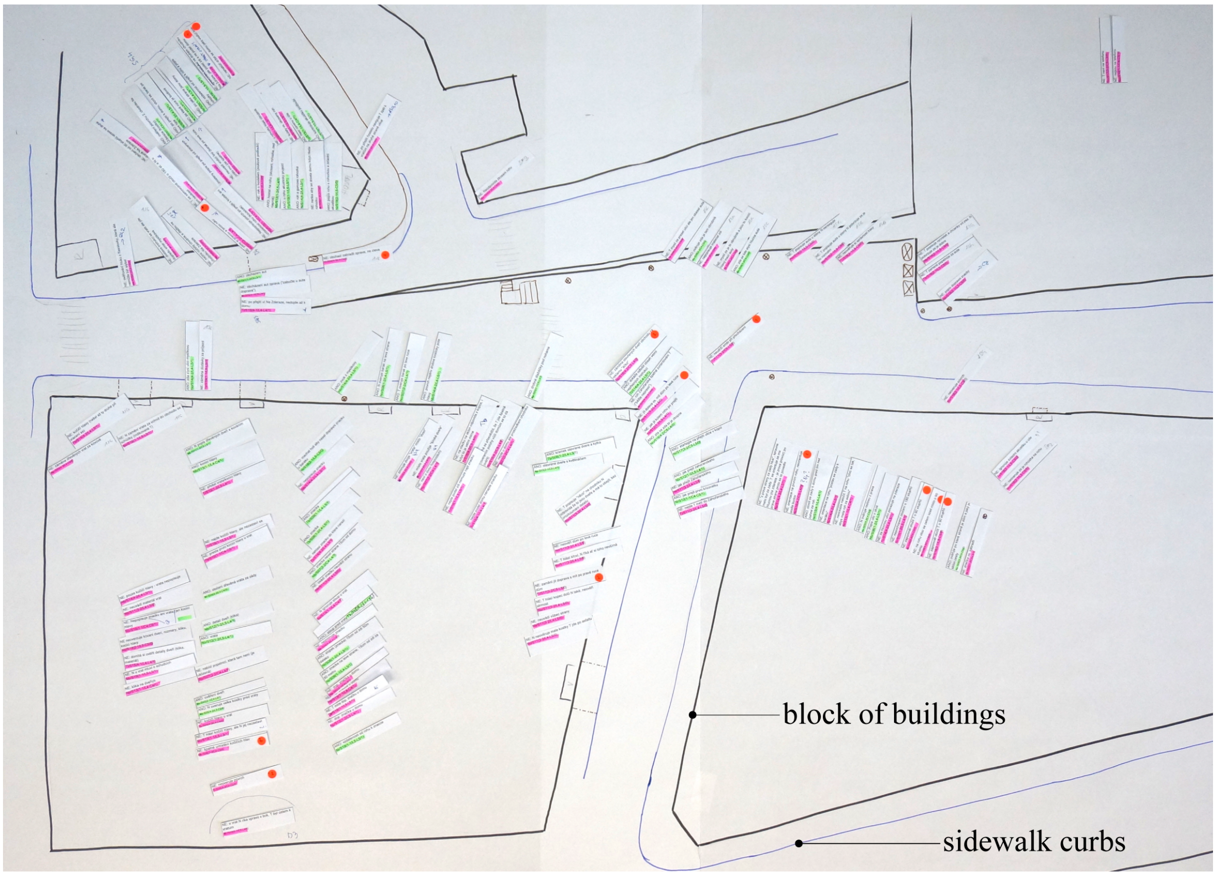


Figure 7.6: Annotated map with data about un/successful navigation strategies used by participants navigating each other in the unknown environment.

we would need to support and tried to generalize the dialog structure to work in all possible situations (see Section 7.3, subsections 7.3.1 and 7.3.2). Next, we mapped the intents to edges traversing from one dialog state to another.

The design process started with low-fidelity prototypes having their formal description contained in state diagrams drawn on a paper and low-fidelity transcripts of the dialog. We run first Wizard of Oz experiments only using state diagram on a paper with an experimenter reading the agent's prompts. Further, an electronic prototype was created to capture one possible conversation path. However, this time using speech to text output for the Wizard. In this phase, we mostly collected more participant utterances to be used for training the natural language understanding module. We struggled to simulate strict computer natural language understanding and participants tend to use longer sentences (especially when experimenter simulated the agent, see lines 23-26 in Excerpt 7.8) and a lot of grounding.

Prototyping is a useful and cheap technique to evaluate the design, even though to use a synthesized voice alternative approaches have to be taken. To do this, one of the useful resource is the VoiceOver[2] feature on iOS or macOS device together with a plain text file.

---

[2] An screenreader present in all Apple products.

Iterating on Implementation

Once the prototyping phase is finished, and implementation is done, another cycle of user testing should be conducted.

We created the dialog model based on the paper drawn state diagram and provided examples for the natural language understanding module first based on expert knowledge and transcribed utterances from low fidelity prototype evaluation. After an iteration with users, we added sample data from experiments (see Excerpt 7.4 and 7.5). Moreover, we identified missing intents and added intents for the repetition of last utterances, going "back" to the previous segment, and recognizing small-talk.

The Future step is to collect data from the users and use them to train the natural language understanding module more precisely. Further, data about task completion can be used to identify problematic places in the dialog structure.

When designing an experiment, it is important to consider the user group and its skills. Given the participants with visual impairments, who usually struggle using a touchscreen for text input (which is needed in situations when dictation does not work), the Wizard of Oz technique can be recommended to shield the participants from the troubles that the conversational user interface on a touchscreen device can cause.

7.4.2 Implementation

We chose IBM Watson Conversation, which is used for identifying user intents, entities (i.e., keywords) and dialog management. As the content is different for each segment of a route, we separate it from the dialog system. For that purpose, we used modified IBM sample chat app[3].

The user interface is based on the modified IBM sample chat app. We focused mainly on accessibility and Apple VoiceOver support. The results support input via Dictation and output using VoiceOver in combination with WAI-ARIA[4] live regions. This approach enables the user to either use voice input or typing in situations with a high level of noise (or low level of social acceptance and understanding).

For natural language generation, we used the server running GIS and routing algorithms[5].

7.5 Conclusion and Future Work

In this chapter, we discuss the design process of creating a conversational agent in a knowledge rich context. We introduced our design process on the use-case of conversational navigation agent for navigation of pedestrians with visual impairments. We showed how we conducted the initial research and how we tackled the integration of different parts of the agent with external GIS while maintaining mixed initiative interaction.

We aim to provide guidance to HCI designers to use a similar design process. One of the other use-cases may be a conversational cookbook agent, which would enable the users to deal with situations when an ingredient is missing ("Can I substitute shallot with onion") or they do not know how to perform an activity ("What does it mean 'medium fire'?").

---

[3] IBM Watson Conversation – https://www.ibm.com/watson/services/conversation/

[4] W3C WAI-ARIA – https://www.w3.org/TR/wai-aria/

[5] Naviterier – http://www.naviterier.cz

Next research steps can lead towards including knowledge structures, which would enable the conversational agent to utilize user's particular preference connected to particular places. Further, we can focus on how to include persuasion strategies and habitual change strategies into the agent.

Speech as an interaction modality can help overcome the limits of WIMP GUI, today's dominant user interface style. Introducing conversational user interfaces can increase accessibility of systems (especially for visually or motor impaired people) or enable interaction in specific contexts (person checking recipe system while cooking and having dirty hands or driver safely operating car infotainment system).

In conversational user interfaces, we operate with user intents which naturally associates objects with actions, what corresponds to task-oriented non-command user interfaces defined by Jakob Nielsen [74]. The task-oriented multimodal UI design approach has a potential to increase safety in situations where primary task is outside the system user interacts with (e.g., driving a car), improve the user experience with setting highly complex system settings, controlling the level of detail of acquired information, or managing the course of mixed-initiative interaction.

Conversational user interfaces have great potential to minimize the motor and visual interaction needed to interact with systems. Thus in the future, we envision embedded conversational user interfaces in many use-cases where the visual or motor interaction is in some way limited. A possible use case can be a blind pedestrian interacting with the city infrastructure through a conversational user interface equipped with single push-to-talk button embedded in a white cane.

7.6 Implications

In this chapter, we showed how to model and prototype a conversational agent for navigation of visually impaired pedestrians. We used findings and results from Chapters 4 and 6 for dialog design and natural language generation for verbal descriptions from GIS. Further research should focus on penalization of routing and implicit level of verbal description level of detail.

In the last chapter, we conclude the thesis, provide the answers to the research questions (see Chapter 1), review the contributions and limitations of the thesis, and provide opinions on future work.

Chapter 8

Conclusion

Abstract. *The last chapter discusses the results in a relation to the research problem and research questions defined in the first chapter. Next, it reflects on the contribution of the thesis and suggests a future research directions.*

8.1 Introduction

In the first chapter, we identified a research problem in how the current navigation systems support the specific needs of visually impaired pedestrians in terms of not providing a verbal description that would support them in gaining sufficient overview and survey knowledge.

Throughout the chapters, we aimed to get insights into the habits and skills of people with visual impairments regarding verbal description (spatial language) of their routes which are suitable sources of survey knowledge to create mental models of the routes by another person [32]. We aimed to design and build prototype systems, which would help utilize the verbal descriptions either while sharing with another person or being automatically generated using a computer system or provided by a conversational agent.

In the next section, we provide an analysis of the answers to the following research questions defined in Chapter 1:

- **RQ1** How do people with visual impairments share spatial knowledge in a form of verbal route description?
- **RQ2** How much and how well do people with visual impairments remember their regular routes?
- **RQ3** How do people with visual impairments perform while sharing the verbal route description over the phone?
- **RQ4** How can be tele-assistance navigation for people with visual impairments improved in general?
- **RQ5** How will people with visual impairments perform with automatically generated verbal description during navigation?

8.2 Discussion

In the Chapters 3, 4 and 6 we provide detailed answers and discussion to the research questions defined in Chapter 1. This section summarizes the answers and discusses them in the context of the thesis and related research.

How do people with visual impairments share spatial knowledge in a form of verbal route description (RQ1)? In the Chapter 3, we showed, that people with visual impairments already share their spatial knowledge with their friends and families. They used a phone call, e-mails or even SMS and they prefer navigation instruction provided by a blind person to those provided by an inexperienced sighted one. These findings are in line with later research on [16], [20], [82], [116]. Further, we quantified the qualitative findings concluding that they often help their friends and family with navigation (53 % several times a year, 10 % several times a month) and they also call to ask for navigation (41 % several times a year, 8 % several times a month), that asking for directions on the street is quite common (47 %) and that they prefer visually impaired navigation (score 3.26 out of 4). We used these findings to further explore the process of tele-assistance between two people with visual impairments (Chapter 4).

How much and how well do people with visual impairments remember their regular routes (RQ2)? Research results summarized in Chapter 3 reveal the following answers. Twenty participants of the study could remember 93.33 km of the regular routes, which makes 4.67 km per person. They mentioned 4.25 destinations on average (restaurants, workplaces, schools, or hobbies) The participants described their regular routes in such detail, we were able to reconstruct precise paths on a map. The description contained both natural and artificial landmarks such as slopes, walls, buildings, or cracks in the pavement. Moreover, interesting patterns emerged such as that congenitally blind participants prefer public transport routes to the walking, however, the sampling was not adjusted for this purpose. The results suggested, that building a collaborative navigation system based on blind-to-blind tele-assistance would be feasible (Chapter 5) and there may be patters on how to select appropriate *navigator* for particular *traveler*.

How do people with visual impairments perform while sharing the verbal route description over the phone (RQ3)? In the Chapter 4, we focused on how people with visual impairment can navigate each other over the phone. If we look at a success rate, 11 out of 18 sessions (1 inconclusive, see Chapter 4) were completed (61.1 % success rate). However, we observed mainly the qualitative aspects of the process – we classified and described the problems and the common situations encountered during navigation in a city. We conclude that the majority of the problems were caused by inaccurate navigation instructions provided by the *navigator*. We also identified how the participants chose strategies when they went astray – the majority of them used backtracking. Most importantly, we observed common strategies for orientation in the streets and avoiding confusion in similar places – paying attention to the position of buildings and slopes. These findings were used to prototype a collaborative navigation system (namely success rate and overall acceptance of the method; see Chapter 5) and to design automatically generated navigation instructions (namely landmarks used to describe the environment and orientation strategies; Chapter 6).

How can be tele-assistance navigation for people with visual impairments improved in general (RQ4)? In the Chapter 4, we provide a set of 5 guidelines – three of them for *navigators* (persons with visual impairments or navigation centers providing tele-assistance) and two of them for *travelers* (pedestrians with visual impairments). The guidelines are based on observed navigation problem and successful strategies leading to reaching the destination. The guidelines are as follows: 1) *Navigator* should describe the environment in detail, with a focus on tactile properties (materials, changes of materials, slopes) and auditory properties (traffic sounds, echoes); 2) *Navigator* should relate the orientation of traveler to the environment, to traveler's previous route, or to auditory landmarks, and navigator should describe landmarks along the route, mention on which hand side is the leading line (building, an edge of a sidewalk), and mention auditory properties; 3) *Navigator* should regularly check *traveler*'s position e.g., ask about execution of instruction and discovered landmarks; 4) *Traveler* should listen to the whole navigation instruction before execution, restate instruction, acknowledge both the understanding and execution of the instruction; 5) *Traveler* should describe the environment in detail, with a focus on tactile and auditory properties. These guidelines, when used, may improve the efficiency of tele-assistance navigation centers and also navigation between friends and family, which was shown to be a regular activity (see Chapter 3).

How will people with visual impairments perform with automatically generated verbal description during navigation (RQ5)? In the chapter 6, we describe several experiments we conducted to evaluate automatically generated verbal descriptions. The error rate with landmark enhanced verbal descriptions compared to those without landmarks (i.e., metric) was comparable (0.16 vs. 0.21 – note that both of the versions used sidewalk based GIS for fair comparison). However, the satisfaction in the terms of comprehension, ambiguity and safety was higher for verbal descriptions enhanced with landmarks. These findings are in line with those of Rehrl et al. [85] who did not find significant difference in error rate or completion time (comparing landmark based vs. metric navigation of sighted pedestrians), however, the difference was in subjective evaluation. In the long-term diary study, we observed success rate of 84.01 % when using a web application prototype participants used for navigation in the wild. Using the same procedure as for automatically generated verbal description we prototyped conversational agent for navigation – verbal descriptions were used for navigation instructions provided by the agent (see Chapter 7).

8.3 Limitations

In this section we discuss the limitation of this thesis both from the methodological and practical point of view.

For the majority of the research questions, we use observational methods. These methods bring insights into the problem and describe how the users interact qualitatively. However, we cannot generalize our findings to the whole population. To be able to claim a measurable profit for the studied population certainly, we would have to conduct large-scale controlled experiments, which are with our group rare (in the context of HCI research field) and hard to organize as several researchers noted [20], [23].

In the chapter 3 we did not analyze quantitatively the occurrence of certain landmarks in the verbal descriptions shared by the participants. Similar data were analyzed by Ross et al. [88] and showed interesting insights into a verbal description of sighted people.

In the chapter 6 we conducted a comparative study in the urban environment. Our motivation was to ensure external validity, i.e., not to evaluate the navigation instructions in the artificial environment as in [93] such as football fields or gyms with artificial walls. We observed real use and interaction. However, the results were compromised (inconclusive) by interruptions from the environment. Furthermore, we do not take the preferred walking speed into account, which may be used to compare the efficiency of navigation systems [28]. However, the walking speed is also influenced by the difficulty of the route, personal experience with independent walking or by use of guiding dog).

The templates used for automatic generation of verbal description are based on a typical layout of European cities with usual block-with-sidewalk type of city centers. This may not apply to cities with different types of layouts or for different cultural settings. For example in countries where in contrast to Europe [91] where people with visual impairments prefer walking and public transport, they use taxis [45] or personal assistants. Moreover, in this thesis, we do not focus on public transportation in the design of automatically generated verbal description of routes (even if it is preferred by the user group). More on how to extend our work with public transport can be found in [11]. Finally, we did not quantify the benefit of the introduced design recommendations.

In the end, GIS used in Chapter 6 is demanding to produce in large scale, and considerable human and financial resources are needed for large-scale deployment. The cost may, however, be partially reduced by use of crowdsourcing of the accessibility attributes [86] without the need of professional on-site reconnaissance (professionals create only “empty” sidewalk network).

8.4 Contributions

To answer the research questions and to propose solutions towards solving the research problem we conducted many studies and designed and evaluated several system prototypes. In this section, we list the major contributions of the thesis.

8.4.1 Practical contributions

We formulated the research problem in the Chapter 1 as follows:

> **Research problem:** We identified a problem in how the current navigation systems support the specific needs of visually impaired pedestrians. Particularly in the way that they do not a provide verbal description that would support them in gaining sufficient route and survey knowledge or the new route.

The contribution of this thesis from the pragmatic research paradigm point of view to solving the research problem we identified is threefold.

First, we proposed a concept collaborative navigation system enhanced by gamification elements, where people (even strangers) with visual impairments directly exchange and narrate verbal descriptions of their regular routes via tele-assistance – this concept is based on their habit of sharing regular routes, remarkable memory of regular routes (Chapter 3) and their performance when being navigated by another person via phone (Chapter 4). In this way, the information needed to form mental repression of a route is being transferred, to the extent of the *navigator's* memory.

Second, we designed, implemented and evaluated the prototype of a navigation system which automatically generates a verbal description of the routes inspired by those made by people with visual impairments (based on sidewalks, crossings, landmarks, slopes or shapes of corners; Chapter 4). By utilizing GIS enhanced with landmarks people with visual impairments naturally use, we aimed to provide a similar experience as if a person is navigated by another person with visual impairment, to the extent limited by the resolution of the GIS.

Third, we prototyped the conversational navigation agent whose behaviour is inspired by that of people with visual impairment (the strategies of proceeding checking or landmark checking) and whose utterances (natural language generation) are automatically generated verbal route descriptions (based on sidewalks, crossings, landmarks, slopes or shapes of corners; Chapter 4). In this way, we simulated the *navigator* using a conversational navigation agent and we gave a hint of future use-cases where using natural language conversation users may change the behavior of the agent regarding the level of information detail provided or properties of the routing (see Chapter 7).

8.4.2 Scientific contributions

The contributions of this work to human-computer interaction and particularly accessibility field may be split into two categories: first, empirical contributions (C1, C2, C4); second, artifact contributions (C3, C5) [117]:

C1 The description of the sharing habits of verbal route descriptions among people with visual impairments including the quantitative analysis and analysis of the quantity of information shared (see Chapter 3).

C2 The observation of people with visual impairments navigating another person with visual impairments via phone with detail description of the situations occurred and successful strategies they took. Further, we provide a set of guidelines to improve tele-assistance of people with visual impairments in general (see Chapter 4).

C3 The concept and design guide for the creation of a serious game for collaborative navigation of people with visual impairments (see Chapter 5).

C4 The formal description of the structure of the automatically generated verbal route description accompanied by the parameters for safe routing and set of design guidelines for future improvements (see Chapter 6).

C5 The description of a design process we took to create a prototype of conversational navigation agent including detailed analysis of dialog management authoring (see Chapter 7).

Future researchers, navigation system designers or conversation analysts may benefit from the descriptions and observations in a way of understating the processes, motivations and strategies behind verbal description sharing. All the concepts and prototypes provide design guidelines and detailed analysis of participant's behaviour and reactions may help to create better user experiences, navigation instructions or more naturally precieved conversational agents.

8.5 Future Work and Final Remarks

Throughout the chapters of this thesis showed possible solutions to the research problem via collaboration, landmark-based navigation, and conversation. We see areas for future development in the sense of adaptation and personalization and situational level of detail through the natural language conversation.

The natural language conversation provides a natural way how people interact with each other in different activities like inquiry, request, story or instructions, compliment, and others [72]. For people with specific needs namely for those with mobility and visual impairments, this type of interaction prevail over other modes of interaction (e-mails, SMS) based on traditional WIMP interfaces. Even from the accessibility point of view, many accessibility solutions enhance WIMP interface with voice command and control or voice output.

Adaptation and Personalization Let us imagine utterances like *"I don't like this crossing," "I always go to the left here,"* or *"I prefer crossings with light signalization,"* may infer how the routing is adjusted on a global level or at a particular spot or ad-hoc during navigation. Providing natural language conversation also avoid cumbersome setting of routing algorithms (for cars, we are often offered toll/without a toll, allow unpaved road, ferry) for a large number of options available when routing takes part on sidewalks and crossings.

Situational level of detail Furthermore, utterances like *"Tell me more about shops around me,"* or *"I don't need so much information on this route,"* may influence how future navigation system adjust the level of detail either spontaneously or deliberately tied to specific locations or routes. We offered first insights into how ad-hoc route changing and level of detail may work in Chapter 7.

Another direction for the future research may be in the adaptation of the automatic verbal description generation (Chapter 6) to new environments like indoor, countryside or artificial worlds, which may be used in the future conversational navigation agents. The leaping enhances of the natural language conversation will have a fundamental impact on accessibility for people with visual or motor impairments – moving the interfaces from UI driven to conversational driven ones.

THE END

Bibliography

[1] C. C. Abt, *Serious games.* University Press of America, 1987.

[2] J. F. Allen, D. K. Byron, M. Dzikovska, G. Ferguson, L. Galescu, and A. Stent, "Toward conversational human-computer interaction", *AI magazine*, vol. 22, no. 4, p. 27, 2001.

[3] J. Allen, C. I. Guinn, and E Horvtz, "Mixed-initiative interaction", *IEEE Intelligent Systems and their Applications*, vol. 14, no. 5, pp. 14–23, 1999.

[4] American Foundation for the Blind. (2018). Giving directions to a person who is blind or visually impaired, [Online]. Available: `{http://www.afb.org/info/friends-and-family/etiquette/giving-directions/235}` (visited on 06/06/2018).

[5] M. Avila, M. Funk, and N. Henze, "Dronenavigator: Using drones for navigating visually impaired persons", ser. ASSETS 2015, Lisbon, Portugal: ACM, 2015, pp. 327–328, ISBN: 978-1-4503-3400-6.

[6] M. Avila, K. Wolf, A. Brock, and N. Henze, "Remote assistance for blind users in daily life: A survey about be my eyes", in *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, ACM, 2016, p. 85.

[7] R. M. Baecker, "Timelines themes in the early history of hci—some unanswered questions", *Interactions*, vol. 15, no. 2, pp. 22–27, 2008.

[8] J Balata, J Franc, Z Mikovec, and P Slavik, "Collaborative navigation of visually impaired", *Journal on Multimodal User Interfaces*, vol. 8, no. 2, pp. 175–185, 2014.

[9] J Balata, Z Mikovec, and I Maly, "Navigation problems in blind-to-blind pedestrians tele-assistance navigation", in *INTERACT*, Springer, 2015, pp. 89–109.

[10] J Balata, Z Mikovec, and P Slavik, "Mutual communication in navigation of visually impaired", *CogInfoCom*, vol. 445, pp. 08 854–4141, 2012.

[11] J. Balata, J. Berka, and Z. Mikovec, "Indoor-outdoor intermodal sidewalk-based navigation instructions for pedestrians with visual impairments", in *Computers Helping People with Special Needs*, K. Miesenberger and G. Kouroupetroglou, Eds., Cham: Springer International Publishing, 2018, pp. 292–301, ISBN: 978-3-319-94274-2.

[12] J. Balata, Z. Mikovec, P. Bures, and E. Mulickova, "Automatically generated landmark-enhanced navigation instructions for blind pedestrians", in *Computer Science and Information Systems (FedCSIS), 2016 Federated Conference on*, IEEE, 2016, pp. 1605–1612.

[13] J. Balata, Z. Mikovec, and P. Slavik, "Conversational agents for physical world navigation", in *Studies in Conversational UX Design*, R. J. Moore, M. H. Szymanski, R. Arar, and G.-J. Ren, Eds. Springer, 2018, pp. 1–23.

[14] ——, "Landmark-enhanced route itineraries for navigation of blind pedestrians in urban environment", *Journal on Multimodal User Interfaces*, 2018, ISSN: 1783-8738. DOI: `10.1007/s12193-018-0263-5`. [Online]. Available: `https://doi.org/10.1007/s12193-018-0263-5`.

[15] J. Balata, Z. Mikovec, P. Slavik, and M. Macik, "Game aspects in collaborative navigation of blind travelers", in *Handbook of Research on Holistic Perspectives in Gamification for Clinical Practice*, IGI Global, 2016, pp. 497–523.

[16] N. Banovic, R. L. Franz, K. N. Truong, J. Mankoff, and A. K. Dey, "Uncovering information needs for independent spatial learning for users who are visually impaired", in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM, 2013, p. 24.

[17] J. P. Bigham, C. Jayant, H. Ji, G. Little, A. Miller, R. C. Miller, R. Miller, A. Tatarowicz, B. White, S. White, *et al.*, "Vizwiz: Nearly real-time answers to visual questions", in *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, ACM, 2010, pp. 333–342.

[18] "Spatial language", in *Encyclopedia of Neuroscience*, M. D. Binder, N. Hirokawa, and U. Windhorst, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 3799–3799, ISBN: 978-3-540-29678-2. DOI: `10.1007/978-3-540-29678-2_5547`. [Online]. Available: `https://doi.org/10.1007/978-3-540-29678-2_5547`.

[19] R. R. Bourne, S. R. Flaxman, T. Braithwaite, M. V. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen, J. Leasher, H. Limburg, *et al.*, "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: A systematic review and meta-analysis", *The Lancet Global Health*, vol. 5, no. 9, e888–e897, 2017.

[20] N. A. Bradley and M. D. Dunlop, "An experimental investigation into wayfinding directions for visually impaired people", *Personal and Ubiquitous Computing*, vol. 9, no. 6, pp. 395–403, 2005. DOI: `10.1007/s00779-005-0350-y`.

[21] A. Brock, "Interactive maps for visually impaired people: Design, usability and spatial cognition", PhD thesis, Université Toulouse 3 Paul Sabatier, 2013.

[22] T. H. Bui, M. Poel, A. Nijholt, and J. Zwiers, "A tractable hybrid ddn–pomdp approach to affective dialogue modeling for probabilistic frame-based dialogue systems", *Natural Language Engineering*, vol. 15, no. 02, pp. 273–307, 2009.

[23] M Bujacz, P Baranski, M Moranski, P Strumillo, and A Materka, "Remote guidance for the blind – a proposed teleassistance system and navigation trials", in *HSI 2008*, IEEE, 2008, pp. 888–892. DOI: `10.1109/hsi.2008.4581561`.

[24] T. J. Carroll, *Blindness: What it is, what it does, and how to live with it.* Little, Brown, 1961.

[25] S. M. Casey, "Cognitive mapping by the blind.", *Journal of Visual Impairment & Blindness*, 1978.

[26] J. Chu-Carroll and M. K. Brown, "Tracking initiative in collaborative dialogue interactions", in *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, 1997, pp. 262–270.

[27] G. Chung, "Developing a flexible spoken dialog system using simulation", in *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, Association for Computational Linguistics, 2004, p. 63.

[28] D. Clark-Carter, A. Heyes, and C. Howarth, "The efficiency and walking speed of visually impaired people", *Ergonomics*, vol. 29, no. 6, pp. 779–789, 1986. DOI: `10.1080/00140138608968314`.

[29] Czech Blind United, *Annual Report 2013*, 2013. [Online]. Available: `{http://archiv.sons.cz/docs/Vyrocni_zprava_SONS_2013.pdf}` (visited on 06/06/2018).

[30] ——, (2018). Sjednocená organizace nevidomých a slabozrakých ČR, [Online]. Available: `{http://www.sons.cz/}` (visited on 06/06/2018).

[31] M. Denis, “Assessing the symbolic distance effect in mental images constructed from verbal descriptions: A study of individual differences in the mental comparison of distances”, *Acta Psychologica*, vol. 127, no. 1, pp. 197–210, 2008.

[32] ——, “The description of routes: A cognitive approach to the production of spatial discourse”, *Cahiers de psychologie cognitive*, vol. 16, no. 4, pp. 409–458, 1997.

[33] M. Denis and M. Zimmere, “Analog properties of cognitive maps constructed from verbal descriptions”, *Psychological Research*, vol. 54, no. 4, pp. 286–298, 1992.

[34] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, “From game design elements to gamefulness: Defining gamification”, in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, ACM, 2011, pp. 9–15.

[35] R. M. Downs, *Image & environment: Cognitive mapping and spatial behavior.* Aldine, 2005.

[36] DPP. (2018). Profil spolecnosti, [Online]. Available: `{http://www.dpp.cz/profil-spolecnosti/}` (visited on 06/06/2018).

[37] P. Dragicevic, *HCI statistics without p-values*, 2015.

[38] M. A. Espinosa, S. Ungar, E. Ochaíta, M. Blades, and C. Spencer, “Comparing methods for introducing blind and visually impaired people to unfamiliar urban environments”, *Journal of Environmental Psychology*, vol. 18, pp. 277–287, 1998.

[39] J. Faria, S. Lopes, H. Fernandes, P. Martins, and J. Barroso, “Electronic white cane for blind people navigation assistance”, in *WAC 2010*, IEEE, 2010, pp. 1–7.

[40] J. F. Fletcher, “Spatial representation in blind children. 1: Development compared to sighted children.”, *Journal of Visual Impairment and Blindness*, vol. 74, no. 10, pp. 381–85, 1980.

[41] D Frohlich and P Luff, “Applying the technology of conversation to the technology for conversation”, *Computers and conversation*, pp. 187–220, 1990.

[42] N. Gale, R. G. Golledge, J. W. Pellegrino, and S. Doherty, “The acquisition and integration of route knowledge in an unfamiliar neighborhood”, *Journal of Environmental Psychology*, vol. 10, no. 1, pp. 3 –25, 1990, ISSN: 0272-4944. DOI: `https://doi.org/10.1016/S0272-4944(05)80021-0`. [Online]. Available: `http://www.sciencedirect.com/science/article/pii/S0272494405800210`.

[43] F. Gaunet, J.-L. Martinez, C. Thinus-Blanc, *et al.*, “Early-blind subjects’ spatial representation of manipulatory space: Exploratory strategies and reaction to change”, *Perception*, vol. 26, pp. 345–366, 1997.

[44] F. Gaunet, C. Thinus-Blanc, *et al.*, “Early-blind subjects’ spatial abilities in the locomotor space: Exploratory strategies and reaction-to-change performance”, *Perception*, vol. 25, pp. 967–982, 1996.

[45] A. J. R. Godfrey and D. M. Brunning, “Reconciling true and incurred costs of blindness in new zealand”, *Social Policy Journal of New Zealand*, no. 36, p. 208, 2009.

[46] R. G. Golledge, “Geography and the disabled: A survey with special reference to vision impaired and blind populations”, *Tran. of the Inst. of British Geographers*, pp. 63–85, 1993. DOI: `10.2307/623069`.

[47] ——, “Human wayfinding and cognitive maps”, *Wayfinding behavior: Cognitive mapping and other spatial processes*, pp. 5–45, 1999.

[48] R. G. Golledge, R. L. Klatzky, and J. M. Loomis, “Cognitive mapping and wayfinding by adults without vision”, in *The construction of cognitive maps*, Springer, 1996, pp. 215–246.

[49] K. Hammond, R. Burke, C. Martin, and S. Lytinen, "Faq finder: A case-based approach to knowledge navigation", in *Proceedings of the 11th Conference on Artificial Intelligence for Applications*, ser. CAIA '95, Washington, DC, USA: IEEE Computer Society, 1995, pp. 80–, ISBN: 0-8186-7070-3.

[50] K. Hara, S. Azenkot, M. Campbell, C. L. Bennett, V. Le, S. Pannella, R. Moore, K. Minckler, R. H. Ng, and J. E. Froehlich, "Improving public transit accessibility for blind riders by crowdsourcing bus stop landmark locations with google street view: An extended analysis", *ACM Transactions on Accessible Computing (TACCESS)*, vol. 6, no. 2, p. 5, 2015.

[51] K. Hara, V. Le, and J. Froehlich, "Combining crowdsourcing and google street view to identify street-level accessibility problems", in *CHI 2013*, ACM, 2013, pp. 631–640.

[52] H. Heft, "The role of environmental features in route-learning: Two exploratory studies of way-finding", *Environmental Psychology and Nonverbal Behavior*, vol. 3, no. 3, pp. 172–185, 1979.

[53] H. Hile, R. Grzeszczuk, A. Liu, R. Vedantham, J. Košecka, and G. Borriello, "Landmark-based pedestrian navigation with enhanced spatial reasoning", in *Pervasive Computing*, Springer, 2009, pp. 59–76.

[54] E. W. Hill, J. J. Rieser, M.-M. Hill, M. Hill, *et al.*, "How persons with visual impairments explore novel spaces: Strategies of good and poor performers.", *Journal of visual impairment & blindness*, 1993.

[55] R. T. Høegh, J. Kjeldskov, M. B. Skov, and J. Stage, "A field laboratory for evaluating in situ", *Handbook of Research on User Interface Design and Evaluation for Mobile Technology, IGI Global*, pp. 982–996, 2008.

[56] Z Hunaiti, V Garaj, W Balachandran, and F Cecelja, "Use of remote vision in navigation of visually impaired pedestrians", in *International Congress Series*, Elsevier, vol. 1282, 2005, pp. 1026–1030.

[57] T. Ishikawa, H. Fujiwara, O. Imai, and A. Okabe, "Wayfinding with a gps-based mobile navigation system: A comparison with maps and direct experience", *Journal of Environmental Psychology*, vol. 28, no. 1, pp. 74–82, 2008. DOI: `10.1016/j.jenvp.2007.09.002`.

[58] ISO, *Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems*, 2015. [Online]. Available: `https://www.iso.org/standard/52075.html` (visited on 06/06/2018).

[59] D. Jurafsky, *Speech & language processing.* Pearson Education India, 2000.

[60] V. A. Kulyukin, J. Nicholson, D. A. Ross, J. R. Marston, and F. Gaunet, "The blind leading the blind: Toward collaborative online route information management by individuals with visual impairments.", in *AAAI Spring Symposium: Social Information Processing*, 2008, pp. 54–59.

[61] G. Leshed, T. Velden, O. Rieger, B. Kot, and P. Sengers, "In-car gps navigation: Engagement with and disengagement from the environment", in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2008, pp. 1675–1684.

[62] J. M. Loomis, R. G. Golledge, and R. L. Klatzky, "Navigation system for the blind: Auditory display modes and guidance", *Presence: Teleoperators and Virtual Environments*, vol. 7, no. 2, pp. 193–203, 1998. DOI: `10.1162/105474698565677`.

[63] J. M. Loomis, R. G. Golledge, R. L. Klatzky, J. M. Speigle, and J. Tietz, "Personal guidance system for the visually impaired", in *Assets 1994*, ACM, 1994, pp. 85–91.

[64] J. M. Loomis, R. L. Klatzky, and N. A. Giudice, "Representing 3d space in working memory: Spatial images from vision, hearing, touch, and language", in *Multisensory imagery*, Springer, 2013, pp. 131–155.

[65] M. C. Lovett, "Problem solving", *Stevens' handbook of experimental psychology*, 2002.

[66] N. Mackenzie and S. Knipe, "Research dilemmas: Paradigms, methods and methodology", *Issues in educational research*, vol. 16, no. 2, pp. 193–205, 2006.

[67] A. J. May, T. Ross, S. H. Bayer, and M. J. Tarkiainen, "Pedestrian navigation aids: Information requirements and design implications", *Personal and Ubiquitous Computing*, vol. 7, no. 6, pp. 331–338, 2003. DOI: `10.1007/s00779-003-0248-5`.

[68] S. Millar, *Space and sense.* Psychology Press, 2008.

[69] ——, *Understanding and representing space: Theory and evidence from studies with blind and sighted children.* Oxford University/Clarendon Press, 1994.

[70] A. Millonig and K. Schechtner, "Developing landmark-based pedestrian-navigation systems", *ITSC 2007*, vol. 8, no. 1, pp. 43–49, 2007. DOI: `10.1109/tits.2006.889439`.

[71] M. Modsching, R. Kramer, and K. ten Hagen, "Field trial on gps accuracy in a medium size city: The influence of built-up", in *3rd workshop on positioning, navigation and communication*, 2006, pp. 209–218.

[72] R. J. Moore, R. Arar, G.-J. Ren, and M. H. Szymanski, "Conversational ux design", in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, ACM, 2017, pp. 492–497.

[73] A. Newell, H. A. Simon, *et al.*, *Human problem solving.* Prentice-Hall Englewood Cliffs, NJ, 1972.

[74] J. Nielsen, "Noncommand user interfaces", *Communications of the ACM*, vol. 36, no. 4, pp. 83–99, 1993.

[75] M. L. Noordzij, S. Zuidhoek, and A. Postma, "The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions", *Cognition*, vol. 100, no. 2, pp. 321–342, 2006.

[76] A. Parush, S. Ahuvia, and I. Erev, "Degradation in spatial knowledge acquisition when using automatic navigation systems", in *International Conference on Spatial Information Theory*, Springer, 2007, pp. 238–254.

[77] P. Pasotti, M. B. van Riemsdijk, and C. M. Jonker, "Representing human habits: Towards a habit support agent", in *Proceedings of the 10th International workshop on Normative Multiagent Systems (NorMAS'16)*, ser. LNCS, To appear., Springer, 2016.

[78] R. Passini, G. Proulx, and C. Rainville, "The spatio-cognitive abilities of the visually impaired population", *Environment and Behavior*, vol. 22, no. 1, pp. 91–118, 1990.

[79] R. Passini and G. Proulx, "Wayfinding without vision an experiment with congenitally totally blind people", *Environment and Behavior*, vol. 20, no. 2, pp. 227–252, 1988.

[80] H. Petrie, V. Johnson, T. Strothotte, R. Michel, A. Raab, and L. Reichert, "User-centred design in the development of a navigational aid for blind travellers", in *INTERACT 1997*, Springer, 1997, pp. 220–227. DOI: `10.1007/978-0-387-35175-9_39`.

[81] M. Pielot, B. Poppinga, W. Heuten, and S. Boll, "A tactile compass for eyes-free pedestrian navigation", in *INTERACT 2011*, Springer, 2011, pp. 640–656. DOI: `10.1007/978-3-642-23771-3_47`.

[82] P.-A. Quinones, T. Greene, R. Yang, and M. Newman, "Supporting visually impaired navigation: A needs-finding study", in *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, ACM, 2011, pp. 1645–1650.

[83] L. Ran, S. Helal, and S. Moore, "Drishti: An integrated indoor/outdoor blind navigation system and service", in *Pervasive Computing and Communications, 2004. PerCom 2004. Proceedings of the Second IEEE Annual Conference on*, IEEE, 2004, pp. 23–30.

[84] N. Raz, E. Striem, G. Pundak, T. Orlov, and E. Zohary, “Superior serial memory in the blind: A case of cognitive compensatory adjustment”, *Current Biology*, vol. 17, no. 13, pp. 1129–1133, 2007.

[85] K. Rehrl, E. Häusler, and S. Leitinger, “Comparing the effectiveness of GPS-enhanced voice guidance for pedestrians with metric-and landmark-based instruction sets”, in *Geographic information science*, Springer, 2010, pp. 189–203. DOI: `10.1007/978-3-642-15300-6_14`.

[86] M. Riganova, J. Balata, and Z. Mikovec, “Crowdsourcing of accessibility attributes on sidewalk-based geodatabase”, in *16th IFIP TC 13 International Conference on Human-Computer Interaction — INTERACT 2017 - Volume 10516*, Berlin, Heidelberg: Springer-Verlag, 2017, pp. 436–440, ISBN: 978-3-319-68058-3. DOI: `10.1007/978-3-319-68059-0_46`. [Online]. Available: `https://doi.org/10.1007/978-3-319-68059-0_46`.

[87] J. Ross, L. Irani, M. S. Silberman, A. Zaldivar, and B. Tomlinson, “Who are the crowdworkers?: Shifting demographics in mechanical turk”, in *CHI ’10 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA ’10, New York, NY, USA: ACM, 2010, pp. 2863–2872, ISBN: 978-1-60558-930-5. DOI: `10.1145/1753846.1753873`. [Online]. Available: `http://doi.acm.org/10.1145/1753846.1753873`.

[88] T. Ross, A. May, and S. Thompson, “The use of landmarks in pedestrian navigation instructions and the effects of context”, in *MobileHCI 2004*, Springer, 2004, pp. 300–304. DOI: `10.1007/978-3-540-28637-0_26`.

[89] H. Sacks and E. A. Schegloff, “Two preferences in the organization of reference to persons in conversation and their interaction”, *Everyday language: Studies in ethnomethodology*, pp. 15–21, 1979.

[90] K. Salen and E. Zimmerman, *Rules of play: Game design fundamentals.* MIT press, 2003.

[91] G. Sammer, T. Uhlmann, W. Unbehaun, A. Millonig, B. Mandl, J. Dangschat, and R. Mayr, “Identification of mobility-impaired persons and analysis of their travel behavior and needs”, *Transportation Research Record: Journal of the Transportation Research Board*, no. 2320, pp. 46–54, 2012.

[92] B. Sawyer, “From cells to cell processors: The integration of health and video games”, *Computer Graphics and Applications, IEEE*, vol. 28, no. 6, pp. 83–85, 2008.

[93] V. R. Schinazi, “Spatial representation and low vision: Two studies on the content, accuracy and utility of mental representations”, in *Int. Congress Series*, Elsevier, vol. 1282, 2005, pp. 1063–1067. DOI: `10.1016/j.ics.2005.05.163`.

[94] A. W. Siegel and S. H. White, “The development of spatial representations of large-scale environments.”, *Adv. in child development and behavior*, vol. 10, p. 9, 1975.

[95] G. Skantze, *Error Handling in Spoken Dialogue Systems-Managing Uncertainty, Grounding and Miscommunication.* Gabriel Skantze, 2007.

[96] C. Spencer, M. Blades, and K. Morsley, “The child in the physical environment”, *N. York: Wiley & Sons*, 1989.

[97] T. Strothotte, S. Fritz, R. Michel, A. Raab, H. Petrie, V. Johnson, L. Reichert, and A. Schalt, “Development of dialogue systems for a mobility aid for blind people: Initial design and usability testing”, in *Proceedings of the Second Annual ACM Conference on Assistive Technologies*, ser. Assets ’96, New York, NY, USA: ACM, 1996, pp. 139–144, ISBN: 0-89791-776-6. DOI: `10.1145/228347.228369`.

[98] M. E. Struiksma, M. L. Noordzij, and A. Postma, “What is the link between language and spatial images? behavioral and neural findings in blind and sighted individuals”, *Acta Psychologica*, vol. 132, no. 2, pp. 145–156, 2009.

[99] M. Swobodzinski and M. Raubal, "An indoor routing algorithm for the blind: Development and comparison to a routing algorithm for the sighted", *International Journal of Geographical Information Science*, vol. 23, no. 10, pp. 1315–1343, 2009.

[100] H. Talkenberg, "Electronic guide dogña technical approach on in-town navigation", in *Rank Prize Funds Symposium on Technology to Assist the Blind and Visually Impaired (Grasmere, Cumbria, England)*, 1996.

[101] The Lighthouse Inc. (1995). Disability estimates, [Online]. Available: `{http://li129-107.members.linode.com/research/statistics-on-vision-impairment/disability-estimates/}` (visited on 06/06/2018).

[102] C. Thinus-Blanc and F. Gaunet, "Representation of space in blind persons: Vision as a spatial sense?", *Psychological bulletin*, vol. 121, no. 1, p. 20, 1997. DOI: `10.1037/0033-2909.121.1.20`.

[103] D. W. Tuttle and N. R. Tuttle, *Self-esteem and adjusting with blindness: The process of responding to life's demands.* Charles C Thomas Publisher, 2004.

[104] Tyfloservis. (2018). Rehabilitace slabozrakých a nevidomých, [Online]. Available: `http://www.tyfloservis.cz/` (visited on 06/06/2018).

[105] S. Ungar, "Cognitive mapping without visual experience", *Cognitive mapping: past, present, and future*, vol. 4, pp. 221–248, 2000.

[106] A. Van Dam, "Post-wimp user interfaces", *Communications of the ACM*, vol. 40, no. 2, pp. 63–67, 1997.

[107] T. Völkel, R. Kühn, and G. Weber, "Mobility impaired pedestrians are not cars: Requirements for the annotation of geographical data", in *International Conference on Computers for Handicapped Persons*, Springer, 2008, pp. 1085–1092.

[108] T. Völkel and G. Weber, "Routecheckr: Personalized multicriteria routing for mobility impaired pedestrians", in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*, ACM, 2008, pp. 185–192.

[109] J. Vystrcil, I. Maly, J. Balata, and Z. Mikovec, "Navigation dialog of blind people: Recovery from getting lost", *EACL 2014*, p. 58, 2014.

[110] J. Vystrčil, Z. Míkovec, and P. Slavík, "Naviterier-indoor navigation system for visually impaired", 2012.

[111] M. Wagener-Wender and K. F. Wender, "Expectations, mental representations, and spatial inferences", *Psychology of Learning and Motivation*, vol. 25, pp. 137–157, 1990.

[112] R. W. White and P. Grant, "Designing a visible city for visually impaired users", *Proc. of the 2009 Int. Conf. on Inclusive Design*, 2009.

[113] WHO, *ICD update and revision platform: change the definition of blindness*, 2009. [Online]. Available: `http://www.who.int/blindness/Change%20the%20Definition%20of%20Blindness.pdf` (visited on 06/06/2018).

[114] W. R. Wiener, R. L. Welsh, and B. B. Blasch, *Foundations of orientation and mobility.* American Foundation for the Blind, 2010, vol. 1.

[115] G. Wilcock, "Wikitalk: A spoken wikipedia-based open-domain knowledge access system", in *Proceedings of the Workshop on Question Answering for Complex Domains*, 2012, pp. 57–70.

[116] M. A. Williams, C. Galbraith, S. K. Kane, and A. Hurst, "Just let the cane hit it: How the blind and sighted see navigation differently", in *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*, ACM, 2014, pp. 217–224.

[117] J. O. Wobbrock, “Seven research contributions in hci”, *studies*, vol. 1, no. 1, pp. 52–80, 2012.

[118] J. O. Wobbrock, S. K. Kane, K. Z. Gajos, S. Harada, and J. Froehlich, “Ability-based design: Concept, principles and examples”, *ACM Transactions on Accessible Computing (TACCESS)*, vol. 3, no. 3, p. 9, 2011.

[119] R. Wycherley and B. Nicklin, “The heart rate of blind and sighted pedestrians on a town route”, *Ergonomics*, vol. 13, no. 2, pp. 181–192, 1970.

[120] S. Young, M. Gasic, B. Thomson, and J. D. Williams, “Pomdp-based statistical spoken dialog systems: A review”, *Proceedings of the IEEE*, vol. 101, no. 5, pp. 1160–1179, 2013.

[121] L. Zeng, R. Kühn, and G. Weber, “Improvement in environmental accessibility via volunteered geographic information: A case study”, *Universal Access in the Information Society*, pp. 1–11, 2016.

[122] L. Zeng and G. Weber, “A pilot study of collaborative accessibility: How blind people find an entrance”, in *MobileHCI ’15*, ACM, 2015, pp. 347–356.

[123] M. Zöllner, S. Huber, H.-C. Jetter, and H. Reiterer, *NAVI – A Proof-of-Concept of a Mobile Navigational Aid for Visually Impaired Based on the Microsoft Kinect*. Springer, 2011.

List of Author's Publications

Following pages present list of the author's publication categorized by relevance to the thesis, publication type together with the author's share and citation number[1].

Thesis Related Publications

In Journals with Impact Factor

- J. Balata, Z. Míkovec, and P. Slavík, "Landmark-enhanced route itineraries for navigation of blind pedestrians in urban environment", *Journal on Multimodal User Interfaces*, vol. 2018, pp. 1–18, 2018,
  Author's share: 65 %; **Citations**: 0; **Citations WoS**: 0.

- J. Balata, J. Franc, Z. Míkovec, *et al.*, "Collaborative navigation of visually impaired", *Journal on Multimodal User Interfaces*, vol. 8, no. 2, pp. 175–185, 2014,
  Author's share: 45 %; **Citations**: 12; **Citations WoS**: 8.

In Journals

- P. Bureš, J. Balata, and E. Mulíčková, "Route4all: A novel approach to pedestrian navigation for people with special needs", *Archives of Transport System Telematics*, vol. 10, no. 1, pp. 9–16, 2017,
  Author's share: 45 %; **Citations**: 0; **Citations WoS**: 0.

In Proceedings Indexed by Web of Science

- J. Balata, Z. Míkovec, P. Bureš, *et al.*, "Automatically generated landmark-enhanced navigation instructions for blind pedestrians", in *Proceedings of the 2016 Federated Conference on Computer Science and Information Systems*, 2016, pp. 1605–1612,
  Author's share: 55 %; **Citations**: 4; **Citations WoS**: 3.

- I. Malý, J. Balata, O. Krejčíř, *et al.*, "Qualitative measures for evaluation of navigation applications for visually impaired", in *6th IEEE Conference on Cognitive Infocommunications CogInfoCom 2015*, 2015, pp. 223–228,
  Author's share: 35 %; **Citations**: 1; **Citations WoS**: 1.

- J. Balata, Z. Míkovec, and I. Malý, "Navigation problems in blind-to-blind pedestrians teleassistance navigation", in *Human-Computer Interaction - INTERACT 2015*, ser. Lecture Notes in Computer Science, 2015, pp. 89–109,
  Author's share: 55 %; **Citations**: 2; **Citations WoS**: 2.

---

[1] Indexation of proceedings in Web of Science and number of citations as of August 2nd 2018. Some of the proceedings will eventually be also indexed by Web of Science later this and next year, as of now, they are in category In Other Proceedings.

- J. Balata, Z. Míkovec, and J. Nováček, "Field study: How blind people communicate while recovering from loss of orientation", in *Cognitive Infocommunications (CogInfoCom), 2013 IEEE 4th International Conference on*, 2013, pp. 313–317,
  Author's share: 60 %; **Citations**: 0; **Citations WoS**: 0.

- J. Balata, Z. Míkovec, and P. Slavík, "Mutual communication in navigation of visually impaired", in *CogInfoCom 2012*, 2012, pp. 769–774,
  Author's share: 60 %; **Citations**: 0; **Citations WoS**: 0.

In Other Proceedings

- J. Balata, Z. Mikovec, and P. Slavik, "Conversational agents for physical world navigation", in *Studies in Conversational UX Design*, R. J. Moore, M. H. Szymanski, R. Arar, *et al.*, Eds. Springer, 2018, pp. 1–23,
  Author's share: 55 %; **Citations**: 0; **Citations WoS**: 0.

- J. Balata, J. Berka, and Z. Mikovec, "Indoor-outdoor intermodal sidewalk-based navigation instructions for pedestrians with visual impairments", in *Computers Helping People with Special Needs*, K. Miesenberger and G. Kouroupetroglou, Eds., 2018,
  Author's share: 35 %; **Citations**: 0; **Citations WoS**: 0.

- V. Gintner, J. Balata, J. Bokšanský, *et al.*, "Improving reverse geocoding: Localization of blind pedestrians using conversational ui", in *8th IEEE International Conference on Cognitive InfoCommunications Proceedings*, 2017, pp. 145–149,
  Author's share: 30 %; **Citations**: 0; **Citations WoS**: 0.

- M. Riganová, J. Balata, and Z. Míkovec, "Crowdsourcing of accessibility attributes on sidewalk-based geodatabase", in *Human-Computer Interaction – INTERACT 2017*, ser. Information Systems and Applications, incl. Internet/Web, and HCI, vol. 10516, 2017, pp. 436–440,
  Author's share: 45 %; **Citations**: 0; **Citations WoS**: 0.

- J. Balata and Z. Míkovec, "Designing conversational navigation for blind pedestrians", in *CHI 2017, Workshop on Conversational UX Design*, 2017,
  Author's share: 50 %; **Citations**: 0; **Citations WoS**: 0.

- J. Balata, Z. Míkovec, P. Slavík, *et al.*, "Game aspects in collaborative navigation of blind travelers", in *Handbook of Research on Holistic Perspectives in Gamification for Clinical Practice*. 2015,
  Author's share: 40 %; **Citations**: 0; **Citations WoS**: 0.

- J. Vystrčil, I. Malý, J. Balata, *et al.*, "Navigation dialog of blind people: Recovery from getting lost", in *Proceedings of the Workshop on Dialogue in Motion (DM)*, 2014, pp. 58–62,
  Author's share: 20 %; **Citations**: 0; **Citations WoS**: 0.

- J. Balata, Z. Míkovec, K. Darebný, *et al.*, "Data structures for landmark-based navigation of blind pedestrians", in *2nd Conference on Mobile and Information Technologies in Medicine Conference Proceedings*, 2014, ISBN: 978-80-01-05637-0,
  Author's share: 50 %; **Citations**: 0; **Citations WoS**: 0.

Thesis Non-related Publications

In Journals with Impact Factor

- J. Balata, Z. Míkovec, and T. Slavíček, "Koalaphone: Touchscreen mobile phone ui for active seniors", *Journal on Multimodal User Interfaces*, vol. 9, no. 4, pp. 263–273, 2015,
  Author's share: 60 %; **Citations**: 5; **Citations WoS**: 3.

In Proceedings Indexed by Web of Science

- A. Kutíková, K. Pražáková, P. Slavík, *et al.*, "Senioři a jejich komunikace se zařízeními v prostředí informačních technologií", in *Stárnutí 2016*, 2016,
  Author's share: 16.67 %; **Citations**: 0; **Citations WoS**: 0.
- J. Balata and Z. Míkovec, "Why visually impaired people resist to adopt touchscreen smartphones in the czech republic?", in *7th IEEE International Conference on Cognitive InfoCommunications*, 2016, pp. 49–54,
  Author's share: 65 %; **Citations**: 1; **Citations WoS**: 0.
- J. Balata, Z. Míkovec, and P. Slavík, "Problems of blind chess players", in *6th IEEE Conference on Cognitive Infocommunications CogInfoCom 2015*, 2015, pp. 179–183,
  Author's share: 50 %; **Citations**: 0; **Citations WoS**: 0.
- T. Slavíček, J. Balata, and Z. Míkovec, "Designing mobile phone interface for active seniors: User study in czech republic", in *CogInfoCom 2014 Proceedings*, 2014, pp. 109–114,
  Author's share: 45 %; **Citations**: 1; **Citations WoS**: 1.
- P. Slavík, Z. Míkovec, M. Macík, *et al.*, "Problémy navigace seniorů v prostoru", in *Stárnutí 2014*, 2014, pp. 120–132,
  Author's share: 25 %; **Citations**: 0; **Citations WoS**: 0.
- J. Balata, M. Macík, and Z. Míkovec, "Heat map thermal display for visually impaired", in *CHI '14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014,
  Author's share: 40 %; **Citations**: 0; **Citations WoS**: 0.

In Other Proceedings

- M. Macík, I. Malý, J. Balata, *et al.*, "How can ICT help the visually impaired older adults in residential care institutions: The everyday needs survey", in *8th IEEE International Conference on Cognitive InfoCommunications Proceedings*, 2017, pp. 157–164,
  Author's share: 20 %; **Citations**: 0; **Citations WoS**: 0.
- J. Szczurková and J. Balata, "Supporting the work of operators of tele-assistance centers for visually impaired people", in *Proceedings of the 20th Bilateral Student Workshop CTU Prague*, 2017, pp. 17–19,
  Author's share: 50 %; **Citations**: 0; **Citations WoS**: 0.
- M. Riganová and J. Balata, "Crowdsourcing as the possible solution to inaccurate navigation systems for handicapped people", in *Proceedings of the 20th Bilateral Student Workshop CTU Prague*, 2017, pp. 13–16,
  Author's share: 50 %; **Citations**: 0; **Citations WoS**: 0.
- V. Gintner and J. Balata, "Accessibility of mobile rich internet applications for visually impaired users", in *Proceedings of the 20th Bilateral Student Workshop CTU Prague*, 2017, pp. 21–23,
  Author's share: 50 %; **Citations**: 0; **Citations WoS**: 0.

- J. Balata, A. Makarič, M. Timr, *et al.*, “Explorations into use-cases of conversation systems in campus environment”, in *Proceedings of the 20th Bilateral Student Workshop CTU Prague*, 2017, pp. 47–49,
  Author's share: 20 %; **Citations**: 0; **Citations WoS**: 0.
- A. Kutíková, J. Balata, and Z. Míkovec, “Explorations into ict usage and behavior in travel related activities of wheelchair users”, in *8th IEEE International Conference on Cognitive InfoCommunications Proceedings*, 2017, pp. 231–236,
  Author's share: 40 %; **Citations**: 0; **Citations WoS**: 0.
- J. Balata, Z. Míkovec, and L. Neoproud, “Blindcamera: Central and golden-ratio composition for blind photographers”, in *Proceedings of the 2015 Multimedia, Interaction, Design and Innovation International Conference on Multimedia, Interaction, Design and Innovation*, 2015, pp. 1–8,
  Author's share: 50 %; **Citations**: 4; **Citations WoS**: 1.
- J. Balata, L. Čmolík, and Z. Míkovec, “On the selection of 2d objects using external labeling”, in *CHI '14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014, pp. 2255–2258,
  Author's share: 43 %; **Citations**: 1; **Citations WoS**: 0.
- J. Balata, K. Pražáková, A. Kutíková, *et al.*, “Quido: Arcade game with thermo-haptic feedback”, in *CogInfoCom 2013*, 2013, pp. 923–924,
  Author's share: 30 %; **Citations**: 1; **Citations WoS**: 1.
- J. Tožička, J. Balata, and Z. Míkovec, “Diverse trajectory planning for uav control displays”, in *AAMAS '13 Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, ser. AAMAS '13, 2013, pp. 1411–1412,
  Author's share: 40 %; **Citations**: 1; **Citations WoS**: 1.
- J. Balata, A. Kutíková, K. Pražáková, *et al.*, “Tangible heat maps for navigation of visually impaired users”, in *CHI 2013 — Explorations in Social Interaction Design*, ser. Workshop: Explorations in Social Interaction Design, 2013,
  Author's share: 60 %; **Citations**: 0; **Citations WoS**: 0.
- J. Balata, M. Macík, and Z. Míkovec, “Context sensitive navigation in hospitals”, in *Proceedings of the MobileMed 2013 conference,*, 2013,
  Author's share: 40 %; **Citations**: 0; **Citations WoS**: 0.
- J. Balata, “A visualization system for usability testing of collaborative tasks”, in *POSTER 2012 - 16th International Student Conference on Electrical Engineering*, 2012, pp. 1–5,
  Author's share: 100 %; **Citations**: 0; **Citations WoS**: 0.