

Adaptive Dialogue Management for a Script Knowledge Based Conversational Assistant



Master Thesis

by

Isabel Feustel

Reviewer: Prof. Dr. Dr.-Ing. Wolfgang Minker
Co-Reviewer: Prof. Dr. Birte Glimm
Supervisor: M.Sc. Juliana Miehle

Institute of Communications Engineering
University of Ulm
March 28, 2019

Abstract

The spoken dialogue systems today already fulfill many requirements, and their human-machine interaction works very well. Nevertheless, the users still adapt their communication style to these system. In order to develop appropriate communication style strategies for a spoken dialogue system, we have investigated the communication style factors elaborateness and directness. Within this work, an indoor navigation spoken dialogue system was developed, which has different communication style strategies and states of knowledge implemented. Moreover, a formula for the adaptation of the system's communication style to the user's communication style was developed. The system navigates the user through the University of Ulm based on scripts. In order to create these scripts, a data collection study was carried out in which 97 participants described the selected routes with the help of videos, which have been recorded beforehand. Finally, the system was evaluated in a user study with 30 participants. In summary, the system performed good, but we could find no significant difference between the investigated communication styles. However, the different knowledge states of the system revealed some significant differences. Overall, we were able to demonstrate the potential of the indoor navigation application and gained new insights for the communication styles and the interaction with a non-omniscient system.

I certify that I have prepared this Master Thesis by my own without any inadmissible outside help.

Ulm, March 28, 2019

(Isabel Feustel)

Contents

List of Figures	ix
List of Tables	xi
1. Introduction	1
1.1. Motivation	1
1.2. Objectives and structure of this work	2
2. Related Work	3
2.1. Communication Styles of Dialogue Systems	3
2.2. Speech Navigation Applications	4
3. Concept	7
3.1. System Requirements	7
3.2. Classification of the user's communication style	9
3.3. Adaptive system communication style	11
4. Data Acquisition for Indoor Navigation Scripts	13
4.1. Description of the study	13
4.2. Script generation	15
5. Implementation	17
5.1. Structure	17
5.2. Dialogue Handling	19
5.2.1. Speech Interface	19
5.2.2. Support Vector Machine	21
5.3. Dialogue Actions and their Communication Style	22
5.4. The Dialogue Manager	23
6. Evaluation	25
6.1. Study design	25
6.2. Execution	26
6.2.1. Required resources	26
6.2.2. Acquisition of volunteers	26
6.2.3. Pilot study	26
6.2.4. Adjustments	26
6.2.5. Main study	27
6.3. Results	28

7. Conclusion	37
A. Additional Material	41
A.1. The Scripts	41
A.1.1. The Small Talk Script	41
A.1.2. The Description Script of the first Route	42
A.1.3. The Description Script of the second Route	45
A.1.4. The Description Script of the third Route	49
A.2. Additional Evaluation Data	54

List of Figures

3.1. Example of a small talk dialogue	8
3.2. Example of a navigation dialogue	10
4.1. The video description tool.	14
4.2. Example Script Snippet	16
5.1. A graphical representation of the program flow.	18
5.2. Example of a Chatito File	20
5.3. The FinishNavigation Dialogue Action as Example for a pre-defined dialogue action.	22
5.4. A graphical representation of the dialogue manager decision process. . . .	24

List of Tables

3.1. List of dialogue actions. The actions, which are used by the system only are marked(*).	7
4.1. The distribution of the participants	15
4.2. Variations of the description style	15
6.1. The distribution of study participants regarding their gender, their frequency of use speech based assistants as well as their age.	27
6.2. The distribution of participants among the combinations of dialogue communication styles and knowledge level with the options full knowledge (Full) and half knowledge (Half).	27
6.3. The results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(f)$ is the mean of the ratings of all dialogues with the fixed communication style strategy, $M(r)$ is the mean of the ratings of all dialogues with the random communication style strategy and $M(a)$ is the mean of the ratings of all dialogues with the adaptive communication style strategy. $D(fa) = M(f) - M(a) $ is the difference between $M(f)$ and $M(a)$, and $\alpha(fa)$ is the statistical significance level using the Mann-Whitney-U-Test. The same applies to $D(ra)$ and $\alpha(ra)$. .	29
6.4. The results of the participants' ratings grouped by the different categories of the questionnaire, whereby the rating scale has been inverted for the negatively formulated statements so that the optimal rating is five for every statement. $M(f)$ is the mean of the ratings of all dialogues with the fixed communication style strategy, $M(r)$ is the mean of the ratings of all dialogues with the random communication style strategy and $M(a)$ is the mean of the ratings of all dialogues with the adaptive communication style strategy.	30
6.5. The distribution of the communication style strategies of the preferred dialogues showing how many study participants preferred each option. . .	31
6.6. The results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(k_f)$ is the mean of the ratings of all dialogues with the system having full knowledge and $M(k_p)$ is the mean of the ratings of all dialogues with the system having partial knowledge. $D = M(k_f) - M(k_p) $ is the difference between $M(k_f)$ and $M(k_p)$, and α is the statistical significance level using the Mann-Whitney-U-Test. The differences that are statistically significant ($\alpha < 0.05$) are marked (*). . .	32

List of Tables

6.7.	The results of the participants' ratings grouped by the different categories of the questionnaire, whereby the rating scale has been inverted for the negatively formulated statements so that the optimal rating is five for every statement. $M(k_f)$ is the mean of the ratings of all dialogues with full knowledge and $M(k_p)$ is the mean of the ratings of all dialogues with partial knowledge, $D = M(k_f) - M(k_p)$	33
6.8.	The results of the participants' ratings grouped by the different categories of the questionnaire, whereby the rating scale has been inverted for the negatively formulated statements so that the optimal rating is five for every statement. $M(k_f, x)$ is the mean of the ratings of all dialogues with full knowledge and the communication style strategy $x(f = fixed, r = random, a = adaptive)$, $D(fa) = M(k_f, f) - M(k_f, a)$ and $D(ra) = M(k_f, r) - M(k_f, a)$ are the differences between the mean values.	34
6.9.	The results of the participants' ratings grouped by the different categories of the questionnaire, whereby the rating scale has been inverted for the negatively formulated statements so that the optimal rating is five for every statement. $M(k_p, x)$ is the mean of the ratings of all dialogues with partial knowledge and the communication style strategy $x(f = fixed, r = random, a = adaptive)$, $D(fa) = M(k_p, f) - M(k_p, a)$ and $D(ra) = M(k_p, r) - M(k_p, a)$ are the differences between the mean values.	36
A.1.	The distribution of the communication styles of the user measured within the small talk dialogue.	54
A.2.	The results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(p)$ is the mean of the ratings of all dialogues were declared preferred and $M(np)$ is the mean of the ratings of all dialogues which were declared as not preferred. $D = M(p) - M(np) $ is the difference between $M(p)$ and $M(np)$, and α is the statistical significance level using the Mann-Whitney-U-Test.	55
A.3.	The results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(k_f, x)$ is the mean of the ratings of all dialogues with full knowledge, where $f = fixed$, $r = random$ and $a = adaptive$. α is the statistical significance level using the Kruskal-Wallis-Test. The differences that are statistically significant ($\alpha < 0.05$) are marked (*).	56
A.4.	The results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(k_f, x)$ is the mean of the ratings of all dialogues with partial knowledge, where $f = fixed$, $r = random$ and $a = adaptive$. α is the statistical significance level using the Kruskal-Wallis-Test.	57

A.5. The results of the participants' ratings grouped by the different categories of the questionnaire, whereby the rating scale has been inverted for the negatively formulated statements so that the optimal rating is five for every statement. $M(k_f, x)$ is the mean of the ratings of all dialogues with full knowledge and the communication style strategy $x(f = fixed, r = random, a = adaptive)$ and $M(k_p, x)$ is the mean of the ratings of all dialogues with partial knowledge and the communication style strategy $x(f = fixed, r = random, a = adaptive)$, $D(f) = M(k_f, f) - M(k_p, f)$, $D(r) = M(k_f, r) - M(k_p, r)$ and $D(a) = M(k_f, a) - M(k_p, a)$ are the differences between the mean values.	58
--	----

1. Introduction

1.1. Motivation

Communication is all around us. We talk to our family, room-mates, colleagues and newly also with systems. These systems can be found on our smartphone, web browser or even at home as Alexa, Google Home or similar systems. But the way we talk to the people next to us we do not talk to a system. Closer, we adapt our linguistic style to these systems. But what happens if a system adapts its linguistic style to our personal one? It was proven that the user attitude towards an agent varies due to the output style of the system [3]. Within this thesis we will examine how an adaptation of two linguistic factors within a dialogue system affects the user satisfaction.

There are many factors which describe the linguistic style of a human being (e.g prosody, wording, emotions, etc.). Since we cannot make use of all these factors, we reduced our definition of a communication style to two features: elaborateness and directness. A sentence is elaborate if more words are used than necessary for the sentences meaning. For example, if you are being asked, "Do you like the summer?" you can respond non-verbose with simply saying "Yes", but you can also answer "Yes, I really like the summer." The statements are the same, in this case, an affirmation of the question, but they differ in their elaborateness. Another answer could be, "I really enjoy the sunshine and love to go out then." This answer is also an affirmation, however, it is an indirect one since there is no explicit "yes" included.

In former studies we could show that the elaborateness and the indirectness of a spoken dialogue system have an effect on the user's satisfaction and each user has its own preference (see [12]). Within the study, different fixed communication styles were assigned to the system. Since the users have different preferences regarding the communication style, we want to investigate, if adapting the system communication style to the user's communication style provides a higher user satisfaction. For this, we compare three different kinds of communication style strategies including an adaptive one, which will adapt to the users communication style. All communication style strategies are explained in the Chapter 3.

For the evaluation we needed to find a suitable use case. In order to receive the user's communication style and to make the user feel the changes of the system, we needed a dialogue sequence with a high amount of dialogue turns. This is why we decided to use a context based on script knowledge.

1. Introduction

A lot of all-day activities use script knowledge (e.g. baking a cake, going for the bus, etc.). We gain the knowledge after doing this for the first time and will adapt and refine our knowledge based on our experiences. This knowledge is called script knowledge, because all actions have an order in which they happen (e.g. first doing the dough and then putting the cake in the oven). Currently, systems cannot refine from user descriptions of script knowledge, but the research is going on. The University of Saarbrücken takes much effort in this topic. For our use case we will just make advantage of the dialogue length, which can be increased when using scripts instead of using common slot filling dialogues (they usually have a maximum length based on the slots). Instead of using baking a cake dialogues, we decided to go for indoor-navigation since we can evaluate this in a real environment. We expect users to speak more natural in a real environment than in a test setting.

Compared to usually navigation, which is already used in a daily bases, indoor navigation is a quite new and interesting topic. It suffers from new challenges like the positioning and new description styles. In using indoor navigation we have the advantage of a straight forward script, since the points of interest are in a fixed order, which is different from e.g a baking a cake scenario. In addition, the navigation offers a longer dialogue flow, which allows the user to connect to the system and us to measure the communication style in more detail. Also, the descriptions for a route offer the possibility to modify them easily. Making a description more elaborate can be easily done in adding more details from the route. E.g. some concise description could be "go left" and the elaborate one could then be "go left at the yellow wall".

1.2. Objectives and structure of this work

In this work, we first show some related work in the field of adaptive dialogue systems, linguistic styles, and indoor navigation applications. We continue with the data collection for the creation of the indoor navigation scripts. After that, we describe the functionality and implementation of our adaptive dialogue system followed by our study setup and the evaluation. Finally, we sum everything up and show some future aspects.

2. Related Work

The field of research in dialogue systems is huge, especially with regard to adaptivity. Within this chapter we will present a small selection of interesting work. First of all, we present works that deal with the communication styles of dialogue systems and the relation to user satisfaction. We then consider work related to adaptivity and conclude this chapter with the investigation of indoor navigation applications.

Berg [2] conducted a study on user expectations for a speech dialogue system. He describes today's systems as very limited and explains that users can adapt to the limited interaction styles and use them efficiently, but he suggests that user satisfaction could increase if there were more human-like systems. "Furthermore, when explicitly offering the possibility to naturally interact with the computer, participants were surprised and stated that this would be much easier than command-oriented control." In conclusion, he discovers that users tend towards human-like systems and generally prefer open questions and full sentences in an interaction with a polite system.

2.1. Communication Styles of Dialogue Systems

If you look at the communication between humans, the diversity it consists of is very large. In addition to linguistic factors, there are other factors such as gestures or facial expressions that influence the users affect. Mairesse et al. [11] offer a huge psychological overview for this topic. They investigated how a linguistic style does affect the user and his personality. Studies show the "empirical evidence that adaption to the conversational partner is also beneficial at the personality level through experiments using hand-crafted utterances designed intuitively to express a particular personality" [11].

Forbes [4] tried to model user satisfaction within a tutoring system. For this purpose, communication related parameters, such as the ratio of the user's words to the system's words, were also used. Furthermore, they tried to track the affective state of the user, which resulted in high training prediction but low test prediction. This indicates the relation between communication style and user satisfaction.

Miehle et al. have already discovered cultural differences in communication [14]. Furthermore, they examined the causes for the different communication styles and the cultural preferences for the factors elaborateness and directness [13]. On the one hand, culture does not influence the preference for directness. On the other hand, with regard to elaborateness a distinction could be observed. Furthermore, it was recognized that not

2. Related Work

only the culture influences the linguistic patterns but also the dialogue domain and other users states and traits have an impact. In addition, the role of the system influences the preference for the system communication style.

Kruijff et al. [8][7] have explored a different style of communication. They examined the interpersonal dimension and compared whether the personal (e.g. "I found 3 albums.") or the impersonal style (e.g. "There are 20 albums.") had an influence on the user's attitude towards the system. They examined whether a fixed communication style is better than an adaptation to the user's style. But they found no significant influence of the system output style on the user attitude. They describe a phenomenon as a disruptive factor, which has already been described by Nass et al. [15], saying that a system that refers to itself with a personal pronoun is perceived as more human than a system that does not.

Another example of adaption in spoken dialogue systems is the application by Litman et al. [9]. The main objective was to increase the task success rate by creating a user model that reflects speech recognition errors and therefore changes the dialogue strategy. The system allows different initiatives (system/user/mixed) and differences between confirmations (no/explicit/implicit). Basically, a system-driven dialogue with explicit inquiry always leads to task success, but this increases the dialogue length considerably and feels less like a human dialogue. Instead, Litman points out that a dialogue initiated by the user with no confirmation is most similar to a natural human-to-human interaction. One aspect that is important to us is that a user-based adaption is not ideal for all applications, since a training unit must always be carried out for each user before the dialog. How we deal with this problem is explained in more detail in Chapter 3).

2.2. Speech Navigation Applications

As already mentioned in Chapter 1, indoor navigation is becoming a popular research area. It is way more complex than outdoor navigation since the indoor environment can vary a lot and the user has more options to walk. For example, while driving a car the user has no other possibility than staying on a straight road, but within a big room he can walk wherever he wants to. Werner and Kessel [19], describe these issues and offer a brief overview of the current work in indoor navigation. They separate the research field into three topics: Positioning, Way-Planning and the User-Interface.

They describe positioning as one of the challenging tasks of indoor navigation since GPS signals used in outdoor navigation cannot be used. Nevertheless, Werner and Kessel show many approaches to get the position of the user using WLAN, sensors, radio technology or RFID.

Within this thesis, the indoor navigation is only created as a prototype using scripts and the positioning is not relevant. However, for the future it might be interesting to carry out the positioning purely by linguistic means. The user describes his environment and the system recognizes the current location on the basis of the semantic information the user provided.

2.2. Speech Navigation Applications

For the Way-Planning Werner and Kessel say: "Way-Planning typically constructs a walkable way out of a list of points-of-interests with some constraints." This is very interesting when it comes to construct a Way-Planning tool. Indoor navigation systems need to have some semantic representation of the points-of-interest and conclude the right way for the user. They further mention various user-interfaces like wireless ear-phones, smartphones or desktop computers.

An already working example for an indoor navigation tool is DRISHTI [16]. It was developed to help impaired people find their way in new environments. Ran et al. point out the difficulty that indoor environments can change rapidly and objects may be close to persons. DRISHTI supports both indoor and outdoor navigation and changes the positioning approaches depending on the mode. It has a voice interface and guides people through the surroundings. It also provides information about possible dangers, the current location and generally more information than needed for unimpaired people.

Hurtig et al. [6] present another navigation application. Although this is not an indoor navigation, the system aims at a human-like interaction and is therefore interesting for this work. The tool uses a multimodal approach to give the user the possibility of multiple levels of communication. It is a mobile navigation application that can be operated by speech or pen input. The system uses speech and a graphical representation using a map for the output. They assume that "[i]n many cases, e.g. giving navigational instructions, verbal explanations may not be the appropriate and most effective way of exchanging information." Furthermore, they come to the conclusion that "there is a need for adaptivity and personalization, allowing the use of different input strategies, with responses tailored according to user preferences." This leads to the assumption that adaptivity in a speech interface, depending on the user profile, can also lead to the promotion of user satisfaction.

Belvin et al. [1] present a prototype of a dialogue system for the use in cars. This prototype simulates the human ability to communicate in order to not expose the driver to high cognitive effort. It allows natural conversations and aims to provide maximum usability for the driver. The initial user inputs are limited to queries for turns and distances during navigation, which will be expanded in the course of the work. Probably the biggest problem so far is the speech recognition of street names, because the variety of street names and their pronunciation is huge.

3. Concept

Within this section we will explain the concept of our adaptive spoken dialogue system. We will first take a look at the system requirements and then continue with our formula to adapt the communication style of the system to the one of the user.

3.1. System Requirements

First of all, the system has a speech interface and thus requires the basic modules for a dialogue system including: speech recognition, speech analysis, dialogue management and text-to-speech generation.

We do not need the semantic information of what the user exactly says, because we are only interested in tracking the communication style and the basic intent. For this, we keep the speech analysis simple having all spoken phrases categorized into dialogue actions (see Table 3.1). The system will then react properly to that categorized input. Since we have no available training data we will remain in using a hard-coded rule behavior for the dialogue management.

Our dialogue system will support three different strategies for its communication styles:

1. A fixed style, which will not change during one dialogue sequence.
2. A random style, which is randomly selected at each turn.
3. An adaptive style, which will adapt to the user's style. The adaptation will be explained in Section 3.3.

Accept	Motivate*
Acknowledge	OfferNextNavPoint
AskWellBeing*	Reject
Declare	Request
Emotional*	RequestNextNavPoint
EndNavigation	RequestRepeat
Goodbye	SimpleApologise*
Greet	Thank

Table 3.1.: List of dialogue actions. The actions, which are used by the system only are marked(*).

3. Concept

System: Hi. How are you?
User: Hi, I'm fine.
System: Awesome, this is great! I feel good. I am now going to ask you some questions so I can get to know you.
User: Yes, that's ok for me.
System: Okay. Do you like sports?
User: Yes, I like sports.
System: Okay. I really like watching sports. How has your day been so far?
User: My day was great so far. I had a nice coffee break. Now I'm looking forward to the study.
System: Okay. My day has been good. I had a lot of good talks today! How do you like the weather?
User: The weather is rather bad. Today is not a good weather.
System: Okay. I like this kind of weather. Do you have plans for the evening?
User: No, no plans for today evening.
System: Okay. I guess I will stay at home and relax. What hobbies do you have?
User: I like skiing very much.
System: Okay. I do a lot of outdoor sports! Do you like animals?
User: Yes of course, I like animals. My favourite animal is the Tiger.
System: Okay. I love animals. Thanks. Bye.

Figure 3.1.: Example of a small talk dialogue

However, the proposed system will have two different dialogue scenarios: a small talk scenario, which is used to classify the users communication style, an indoor-navigation dialogue, which is the main application.

The small talk dialogue is completely system-driven as can be seen in Figure 3.1. This means that the system is going to ask the user multiple small talk questions (e.g about the weather, hobbies, plans for the evening, etc.) to classify the communication style of the user. To make the dialogue more natural, the system will give some responses related to previously asked question. We will use question-answer pairs as used by Sakai et al.[18]. They have used these pairs to get a more natural dialogue flow and to have a good transition to their argumentative dialogue.

Since indoor-navigation is a complex task and we want to focus on the communication style, we imitate the behavior with scripts to simplify the use case. All scripts contain the route descriptions, which have a pre-scripted text for each communication style. Instead of using positioning tools to detect if the user is on the right position for the next prompt, the user needs to initiate the next route description by asking the system. The navigation dialogue can be either system or user driven depending on the systems knowledge (see Figure 3.2).

There are three knowledge levels for the system: no, little and full knowledge. With no

knowledge the system will ask the user to describe the full route. Having little knowledge the system will start with describing the route and stop at a given point and ask the user to continue describing. If the system has full knowledge, it will navigate the user till the end of the route. We use these different knowledge levels to get a different kind of interaction with the user, and maybe also a difference in the communication style.

As an optional feature the system can detect positive and negative statements of the user to give simple motivations and some sort of compassion for the user.

3.2. Classification of the user's communication style

Before we explain the adaptive communication style in more detail, we take a look at the classification of elaborateness and directness. Within a previous project, different classification methods were examined to determine the elaborateness and directness. The classifiers such as a Support Vector Machine or neural network were trained with the corpus of the KRISTINA¹ project.

The training dataset contains German dialogues partitioned in dialogue actions which describe the meaning of the statement (e.g accept, declare, request). Overall we had 4887 dialogue actions for the training. Each dialogue action has an utterance containing the spoken text, a participant (user or system), some flags that are not relevant for this thesis, and the values for the elaborateness and the directness. These values have been annotated by students. The scales for each value are from one to five (1=concise/direct and 5=elaborate/indirect).

After further investigations, the Support Vector Machine(SVM) was finally selected as a suitable classifier. In order to determine the elaborateness, the dialogue action and the number of spoken words were used as input features. This results in a unweighted average recall of 0.849, when reducing the scale to 1 (concise) to 3 (elaborate, contains annotated values 3,4,5). In addition to the previously mentioned factors, POS features were used for the classification of the directness. They describe the grammar of a sentence such as subject, verb, preposition, etc. With this additional factor a unweighted average recall of 0.80 could be achieved, when reducing the scale to 1 (direct) to 2 (indirect, contains annotated values 2,3,4,5).

In order to generalize the classifiers and to use them for our project, individual dialogue actions were combined. For example, a MorningGreet and an AfternoonGreet each became a Greet. Since we only distinguish between two values of elaborateness within this project (concise/elaborate), the values 2 and 3 are summarized as elaborate.

¹<http://kristina-project.eu/en/>

3. Concept

System: Hello, nice to meet you. I am an indoor-navigation system. How may i help you?

User: I would like to get to the BECI.

System: Go straight and turn left near the stand containing magazines in order to reach the stairs.

User: OK. I'm at the stairs. What would I do next?

System: The path continues down the stairs.

User: OK. What's next?

System: After reaching the bottom of the stairs, enter the door and turn right.

User: OK. What's next?

System: There is a ramp at the other end of the room.

User: I'm at the room. What do I do next?

System: Turn right at the first cross

User: I've turned right. Where to continue?

System: your way leads you to the right at the end of the corridor

User: I've turned right and there are two doors which one do I take?

System: Sorry, I did not understand that.

User: Where do have to go next?

System: I am very sorry, but I can't help you further. Unfortunately, my knowledge is still incomplete. Can you continue walking and describe the remaining way, please?

User: Ok. I'm walking through the blue door next.

System: Okay, i see.

User: Then I continued on the corridor and at the end of it I turn left.

System: I'm interested in what's next.

User: I'm walking through a plastic curtain and shortly after that I turn right

System: Okay.

User: Then I'm walking through glass door.

System: Could you please repeat that?

User: I walk through a glass door.

System: How to go on?

User: I continue straight ahead and go through a glass door.

System: Okay, i see.

User: I turn right at the stairs and go through glass door

System: I'd like to know where to go.

User: So to the left there's a blue door and that is the BECI door is.

System: Okay, i see.

User: We're finished now.

System: Thank you very much. Have a nice day.

Figure 3.2.: Example of a navigation dialogue

3.3. Adaptive system communication style

As described in the last section, we want to adapt the system's communication style to the user's communication style. For this, we created a formula to slowly adapt the systems output style to the one of the user.

First of all, we create a numeric mapping of our features elaborateness and directness:

$$\begin{aligned} direct &= 1, indirect = 0 \\ elaborate &= 1, concise = 0 \end{aligned}$$

For each of these features we create a communication style $CS(t)$, with CS_u for the user and CS_s for the system and t being the turn number. The initialized values can be seen in Equation 3.1.

For the initial communication style of the user, we extract the elaborateness and directness of each turn from the support vector machine within the small talk dialogue and take the median of it. As initial values for the systems style we set concise and direct.

$$\begin{aligned} CS_u(smalltalk) &= median\{svm_1, svm_2, svm_3, svm_4, svm_5\} \\ CS_u(0) &= CS_u(smalltalk) \\ directness : CS_s(0) &= 1 \\ elaborateness : CS_s(0) &= 0 \end{aligned} \tag{3.1}$$

The following formula (3.2) describes the calculation of the median of the individual values. For $t < 5$, the median is calculated from the existing SVM values and the last communication style value. For $t \geq 5$, the last five SVM values and the last communication style value are taken for the median calculation.

$$M(t) = \begin{cases} median\{svm_1, \dots, svm_t, CS(0)\}, & 0 < t < 5 \\ median\{svm_{t-4}, \dots, svm_t, CS(t-5)\}, & t \geq 5 \end{cases} \tag{3.2}$$

For the communication style at time t , the median is used as long as this value corresponds to either 0 or 1. Otherwise, the last communication style is used, what can be seen in Equation 3.3.

$$CS_{us}(t) = \begin{cases} CS_{us}(t-1) & M(t) = 0.5 \\ M(t) & else \end{cases}, \forall t \in \mathbb{N} \tag{3.3}$$

4. Data Acquisition for Indoor Navigation Scripts

Our indoor navigation dialogue system requires route description scripts. In order to create these scripts, we collected route descriptions within a user study. The setup of the study and the script generation will be explained within this chapter.

4.1. Description of the study

Within this section, we describe our study, including the aim, the setup and the results.

Aim of the study The aim of this study is to acquire data for creating scripts which can be used in an indoor navigation environment. Moreover, we want investigate the different description styles of human users in order to integrate them in our dialogue system.

Setup of the study The task for this study is to describe the routes which are shown in a video. First of all, we chose three routes within the University of Ulm and made videos while walking along the way. All of the routes contained different points of interest and path elements like the cafeteria, an elevator or stairs. The start and end points of the paths are linked so that all paths can be taken one after the other. In other words, route one ends where route two starts and the end point of route two is the starting point of route three. For the videos we used a Panasonic HDC-HS700 with a hand-held stabilizing tool. Small changes (e.g. stutters, applying music) were edited with VSDC Free Video Editor¹.

In order to perform the study, we created a web page with Vue.js² containing a short demographical questionnaire (age, gender and country of residence) and the video description tool shown in Figure 4.1. The tool consists of two parts: the video and the description box. Within the description box, the user has to enter his route descriptions while watching the video. While typing, the video stops automatically and when pressing enter, which adds an item to the list, the video continues. All items can be edited and deleted with a click on the corresponding icons. For preventing misuse, we added a minimum amount of items needed for a video to proceed in combination with a timer which secures that the video has been watched.

In order to find participants, we used the platform clickworker³ since it allows us to

¹<http://www.videosoftdev.com/de/free-video-editor>

²<https://vuejs.org>

³<https://www.clickworker.com>

4. Data Acquisition for Indoor Navigation Scripts

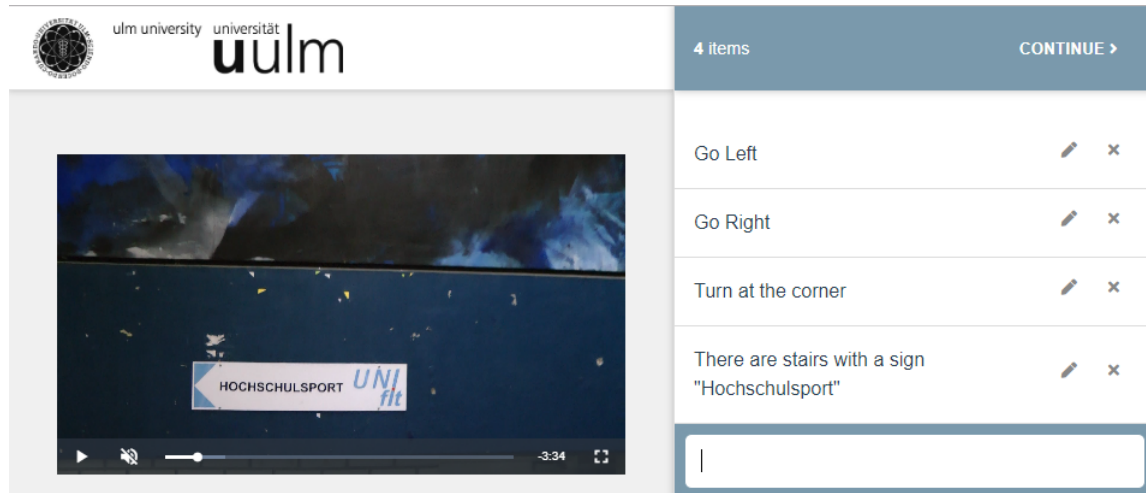


Figure 4.1.: The video description tool.

contact participants with English as their mother tongue. We planned to have 100 participants, including 50 people from the US and 50 from the UK with 25 female and 25 male each.

Before starting the study, we had three German test users who checked if everything works as expected and if the study is easy to understand. The data of two of them is also stored in our database.

Study results Overall we had 97 participants from clickworker with 53 from the UK (28 female, 25 male) and 44 participants from the US (18 male, 25 female, 1 prefer not to tell). Unfortunately, some of the participants misused our tool and did not fill in real data. Moreover, some participants did not finish their descriptions. Table 4.1 shows the distribution of participants in relation to the quality of their data. The column "complete" contains all correct data that can be used after a few small changes (e.g. merging some actions because an unnecessary new item has been created). Some other participants gave correct information, but did not finish the description. We suspect that participants did not see a need to continue the study after the timer expired. Nevertheless, this data can still be used because they it is basically correct although it is incomplete. Any invalid data or data that requires a lot of preparation can be found in the column "wrong/rearrange".

Table 4.2 shows the differences between the individual user groups in terms of the number of descriptions and the number of words per description. There are slight differences between nationality and gender. Finally, we obtained a high variation in the descriptions, which can be used for the generation of our navigation scripts.

	complete	incomplete	wrong/rearrange	total
UK	41	3	9	53
US	33	5	6	44
male	34	3	8	43
female	42	5	6	53
prefer not to tell	0	0	1	1
total	74	8	15	97

Table 4.1.: The distribution of the participants

(a) Average amount of descriptions:					
	overall	uk	us	female	male
video 1	22.50	23.76	20.94	21.76	23.47
video 2	29.72	30.76	28.42	28.95	30.72
video 3	29.92	30.37	29.36	28.83	31.34
(b) Average amount of words per description:					
	overall	uk	us	female	male
video 1	5.79	5.94	5.60	5.92	5.62
video 2	6.03	5.92	6.17	6.19	5.82
video 3	6.23	6.23	6.24	6.28	6.18

Table 4.2.: Variations of the description style

4.2. Script generation

Regneri, Koller and Pinkal[17] have shown a way to automatically generate script knowledge from data. Within a project at the university of Ulm, a student tried to re-implement that Toolkit, so we could use it within this thesis. Unfortunately, we were unable to reproduce the original results. Thus, we decided to do hand-crafted scripts for each route within this thesis.

In order to do so, we reviewed the data collected within the study and selected various descriptions trying to match our communication styles elaborateness and directness. At the end, we created three scripts which look like the example script shown in Figure 4.2.

4. Data Acquisition for Indoor Navigation Scripts

```
<path>
  <pathelement id='1' nextElement='2'>
    <concise_direct>Go straight to the stairs.</concise_direct>
    <elaborate_direct>Go straight and turn left near the stand
      containing magazines in order to reach the stairs.
    </elaborate_direct>
    <concise_indirect>Find the stairs.</concise_indirect>
    <elaborate_indirect>Find the stairs to the left of the stand
      containing magazines. </elaborate_indirect>
  </pathelement>
  <pathelement id='2' nextElement='3'>
    <concise_direct>Go down stairs</concise_direct>
    <elaborate_direct>Go down the stairs and pass the sign saying "
      Hochschulsport".</elaborate_direct>
    <concise_indirect>The path continues down the stairs.
    </concise_indirect>
    <elaborate_indirect>The path continues down the stairs, passing a
      sign saying "Hochschulsport".</elaborate_indirect>
  </pathelement>
  <pathelement id='3' nextElement='4'>
    <concise_direct>Enter the door and turn right</concise_direct>
    <elaborate_direct>After reaching the bottom of the stairs, enter
      the door and turn right.</elaborate_direct>
    <concise_indirect>you can see a door.</concise_indirect>
    <elaborate_indirect>After reaching the bottom of the stairs, you
      can see a door</elaborate_indirect>
  </pathelement>
  ...
</path>
```

Figure 4.2.: Example Script Snippet

5. Implementation

Within this chapter the implementation is explained. First of all, we show the program flow, the package structure and we describe the development environment. After that, we take a closer look the single components of the system.

5.1. Structure

The program consists of two main components: the Dialogue Handler and the Dialogue Manager. The Dialogue Handler manages the general dialogue flow and mainly takes care of the speech interface. The Dialogue Manager decides on the next system actions. The complete program flow is shown in Figure 5.1.

The Dialogue Handler initiates the dialogue with a request for the first system action to the Dialogue Manager. The Dialogue Manager has a stack of system actions which was set within the initialization. The stack is filled with the actions extracted from the XML script depending on the dialogue setting (e.g. SmallTalk.xml or Path1.xml). After the request from the Dialogue Handler, the Dialogue Manager will pop the system action from the stack and returns it to the Dialogue Handler. Within the Dialogue Handler, the system action text is then prompted to Speech. In parallel, a job is running in the background that listens to a possible user input and thus enables barge ins. If a signal is detected, the voice gets recorded and the audio file is sent to the Google Speech-To-Text API. The API then returns a text which is first passed to the RASA Natural Language Understanding(NLU) component. Here the intention of the user is classified from the text and then assigned to a user dialogue action (see Table 3.1 for the dialogue actions). If the classification fails, the Dialogue Handler immediately requests a "RequestRepeat" action from the Dialogue Manager to ask the user for a second input. Otherwise, the user action is enriched by the result of the Support Vector Machine(SVM). The SVM classifies the user's communication style(CS_u) based on the text and the user's intention. The created user action is then passed to the Dialogue Manager with the request for the next system action. The Dialogue Manager then decides whether the normal dialogue flow should continue (pop an action), whether a new action should be created based on the user input (e.g. a repeat) or whether the dialogue should be terminated.

The program is divided into the following four packages:

- **classifier:** This packages includes all necessary files for the support vector machine classification.
- **dialogue:** Contains the main classes for dialogue handling and managing and also

5. Implementation

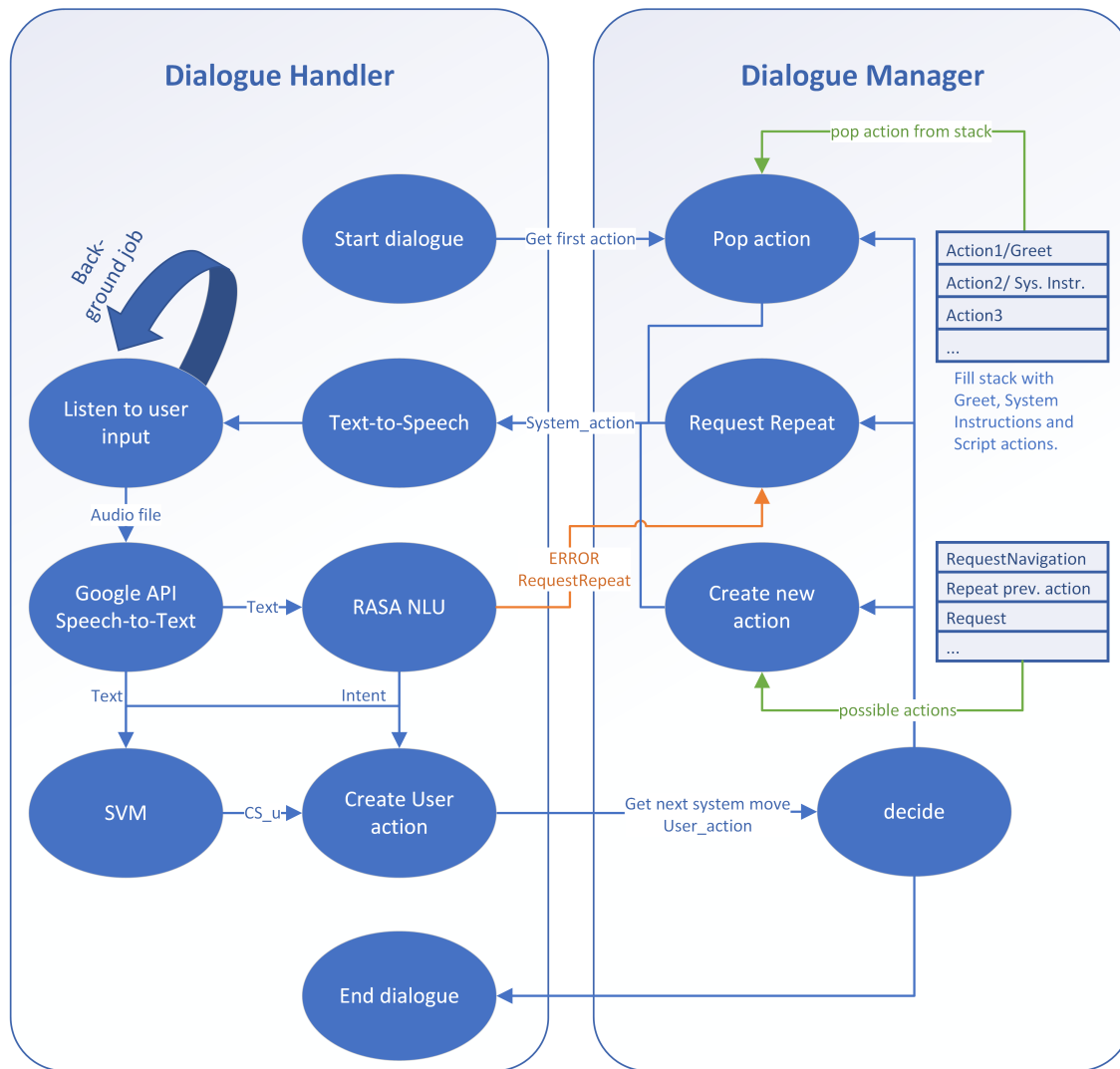


Figure 5.1.: A graphical representation of the program flow.

all predefined dialogue actions and the dialogue action types.

- **nlu:** The interpreter for the language understanding is located within this folder. Also the training data and Chatito(see ??) files are stored here.
- **scripts:** This package contains all XML files containing the small talk question and answer pairs and the route descriptions.

The system was developed with PyCharm and uses Python 3.5 as source programming language. We decided to use Python, because there exist already many libraries to support the implementation of a speech interfaces. Furthermore, the project for the classification of the user communication style was also based on Python and therefore it is easy to import the trained Support Vector Machine(SVM). Moreover, the KRISTINA project has already shown, that the adoption of SVM in other programming languages is possible but more complex and error-prone.

5.2. Dialogue Handling

This section describes the components of the `dialogue_handler.py` class. We will first take a closer look at the parts of the speech interface containing the speech recognition, language analysis and text-to-speech generation. We will then describe the setup of the support vector machine.

5.2.1. Speech Interface

The speech interface can be divided into the three components speech recognition, language analysis and text-to-speech generation. Considering the speech recognition, on initialization we start a background job that listens to the user. Whenever the user speaks to the system the background job will call a callback function and the input will then be recognized by the Google Speech Recognition API, which will return the spoken text. We use the `SpeechRecognition`¹ package based on PyAudio since it can be integrated easy.

For the natural language analysis we decided to use Rasa NLU², which is an open-source natural language processing tool for intent classification and entity extraction in chatbots. It uses tensorflow to train a model that recognizes the intentions and entities of texts based on a given data set. To create that dataset we used Chatito³, which makes the creation of sample sentences and formulations much easier and even gives the possibility to evaluate the classification.

Since many formulations for an intention are very similar and often only single words are exchanged, Chatito offers the possibility to create sentence structures and fill them

¹<https://pypi.org/project/SpeechRecognition/>

²<https://rasa.com/docs/nlu/>

³<https://rodrigopivi.github.io/Chatito/>

5. Implementation

```
1  %[offernextnavpoint]('training': '140', 'testing': '50')
2  ~[description]
3
4  ~[direction]
5  straight
6  down
7  forward
8  ~[sides]
9
10 ~[sides]
11 left
12 right
13
14 ~[color]
15 blue
16 red
17 green
18
19 ~[description]
20 Go through ~[the?] ~[color?] door
21 Go ~[direction]
22 Go ~[direction] till ~[the?] ~[location]
23 Turn ~[sides]
24 Turn ~[sides] at ~[the?] ~[location]
25 Keep going ~[direction]
26 Keep going ~[direction] till ~[the?] ~[location]
27 Take a ~[sides]
28
29 ~[location]
30 hallway
31 stairs
32 road
33 door
34 blue sofa
35 sitting area
36 corridor ends
```

Figure 5.2.: Example of a Chatito File

with words. Figure 5.2 shows an example of a Chatito file. The first line of code shows the intent "offernextnavpoint" and amount of training and testing data. Line 2 is the description for this intent, in this case "[description]", which is a link to the sentences of line 19 and below. The syntax ~[x] can be seen as variable or command. In line 20 for example ~[the?] means that "the" can be optional, whereby ~[color?] is an optional link to the color texts from line 14-17. In total, these 36 lines of script create 190 example sentences, which will be stored in a Rasa NLU syntax json file and can be used for the training of the model. This demonstrates the efficiency of using Chatito. Despite the fact that Chatito facilitates the creation of data, it is an immensely time-consuming and difficult task, since the formulations must be well planned. Furthermore, the neural network requires a lot of data to provide a good detection rate. Overall, we used 2698 generated sentence samples and 10 distinct intents to train the NLU model.

Once the model has been trained, the spoken text of the user can be classified within the Dialogue Manager with a simple call.

For the optional feature of the sentimental analysis we implemented the *SentimentIntensityAnalyzer* from the NLTK⁴ package. The sentiment analyser classifies the user's statements positively or negatively. Within the current system, the dialogue manager can react to the user's mood with regard to the question of well-being and answers appropriately to the mood. If the user is doing well, so the sentiment analysis delivers a positive result, the system is pleased with a positive statement ("Awesome!"). If the user feels bad, so the sentiment analysis delivers a negative result, the system reacts with compassion and says "I hope you feel better soon". It was planned to extend this generally to small talk dialogue to achieve a natural interaction. However, due to other priorities the implementation was not finished.

The speech synthesis was implemented with the package PyTTSx3⁵. The package is platform independent and uses the speech synthesis of the underlying operating system, which in our case was Windows 10. However, the system has not been tested on any other operating system, but we wanted to leave the possibility open.

5.2.2. Support Vector Machine

For the integration of the support vector machine we used the package sklearn⁶. The classifier itself was saved within a pickle file and can be easily loaded with the *joblib* command. However, we implemented three methods to prepare the features vector to fit the classifier input dimension.

One method calculates the amount of words from the user's spoken text and another uses polyglot⁷ to extract the POS-features, which are needed for the classification of the directness. Since the dialogue actions of our system differ from the dialogue action feature

⁴<https://www.nltk.org/>

⁵<https://pypi.org/project/pyttsx3/>

⁶<https://scikit-learn.org/stable/>

⁷<https://pypi.org/project/polyglot/>

5. Implementation

```
1         from dialogue.actions.dialogue_action import DialogueAction
2         from dialogue.actions.dialogue_type import DialogueType
3
4         class FinishNavigation(DialogueAction):
5             action_type = DialogueType.EndNavigation
6
7             sentences = {
8                 'concise_direct': "We've arrived.",
9                 'elaborate_direct': "According to my information, we have
10                    arrived.",
11                 'concise_indirect': "This is your destination.",
12                 'elaborate_indirect': "This is your destination. I don't have
13                    any more directions for you. "
14             }
15
16         def __init__(self):
17             super().__init__(self.action_type, self.sentences)
```

Figure 5.3.: The FinishNavigation Dialogue Action as Example for a pre-defined dialogue action.

of the SVM, we needed to merge some of them. *RequestNextNavPoint* and *RequestRepeat* became a *Request* and *OfferNextNavPoint* and *EndNavigation* became a *Declare*. Besides, the dialogue actions had to be one hot encoded, which we implemented. Finally the input vector had to be scaled. For this, we exported the scaler from the original project and loaded it into our project.

5.3. Dialogue Actions and their Communication Style

Within this section we want to show the implementation of the dialogue actions and the communication styles. We have defined an enum class *dialogue_type.py*, which contains the definitions for all possible dialogue action types. Any dialogue action will derive from the parent class *DialogueAction*, which offers the method *get_sentence(style)* to return the text of the dialogue action in the given communication style. Figure 5.3 shows an example of a dialogue action, which is set as a *EndNavigation* (line 5) and has a dictionary ("sentences" - line 7) containing the texts for the communication styles. Overall, we have seven pre-defined dialogue actions containing fixed texts for each communication styles. These actions are: *Acknowledge*, *AskForAssistance*, *AskWellBeing*, *FinishNavigation*, *Goodbye*, *Greeting*, *RequestNextNavPoint*, *RequestRepeat*, *Thanks*.

RequestNextNavPoint is an exception because it contains not only one text-dictionary with the communication styles but multiple rephrases for the action (e.g. "Where to

go?", "What's next?", etc). This means that the variable *sentences* is a list of dictionaries instead of a single dictionary and the *get_sentence(style)* method of the parent class is reimplemented within this class to apply the selection to a list of dictionaries.

5.4. The Dialogue Manager

The main component of our system is the dialogue manager where our designed formula for the adaptive system communication style is applied.

Within the initialization of the Dialogue Manager a queue is filled with newly created actions extracted from the scripts, which represent the dialogue sequence. If the knowledge of the system is set to partial, it will only extract some descriptions of the script. Besides, it is possible to set the default systems communication style, which is used as starting point and persistent value for the fixed communication style strategy. The Dialogue Manager has a method *set_communication_style()*, which sets the communication style based on its strategy settings(fixed, random, or adaptive).

Figure 5.4 shows the whole decision process of the Dialogue Manager. When the Dialogue Handler requests the next system action the process is triggered. First, the system's communication style is updated based on the communication style strategy. Afterwards, some conditions are verified to choose the next step. If the incoming user action is a *RequestRepeat*, the system will repeat the last system action and if the user action is stated as *Goodbye* the system will confirm the termination of the dialogue with an explicit request to the user. When the user confirms the previous request for ending the dialogue, the system will terminate with a *Thank* and a *Goodbye* action.

As long as there are dialogue actions in the queue, the system will pass the actions on to the user. The Smalltalk dialogue is finished when there are no more entries in the queue and the system will return a *Thank* and *Goodbye* action to the user. However, for the navigation dialogue, more conditions have to be met until the dialogue ends. First, the system will check if the user describes the route and if he has finished the Navigation (*EndNavigation*). If this is the case, the dialogue is ended again with a *Thank* and a *Goodbye*. The same applies if the system has full knowledge, because it means that at this point there are no further actions in the queue and thus the system will finish with an *EndNavigation* followed by a *Thanks* and a *Goodbye*. If the system uses partial knowledge, it will ask the user for assistance if not done beforehand. Otherwise, the system will respond to the users description in alternating order with an *Acknowledge* or *RequestNextNavPoint*.

5. Implementation

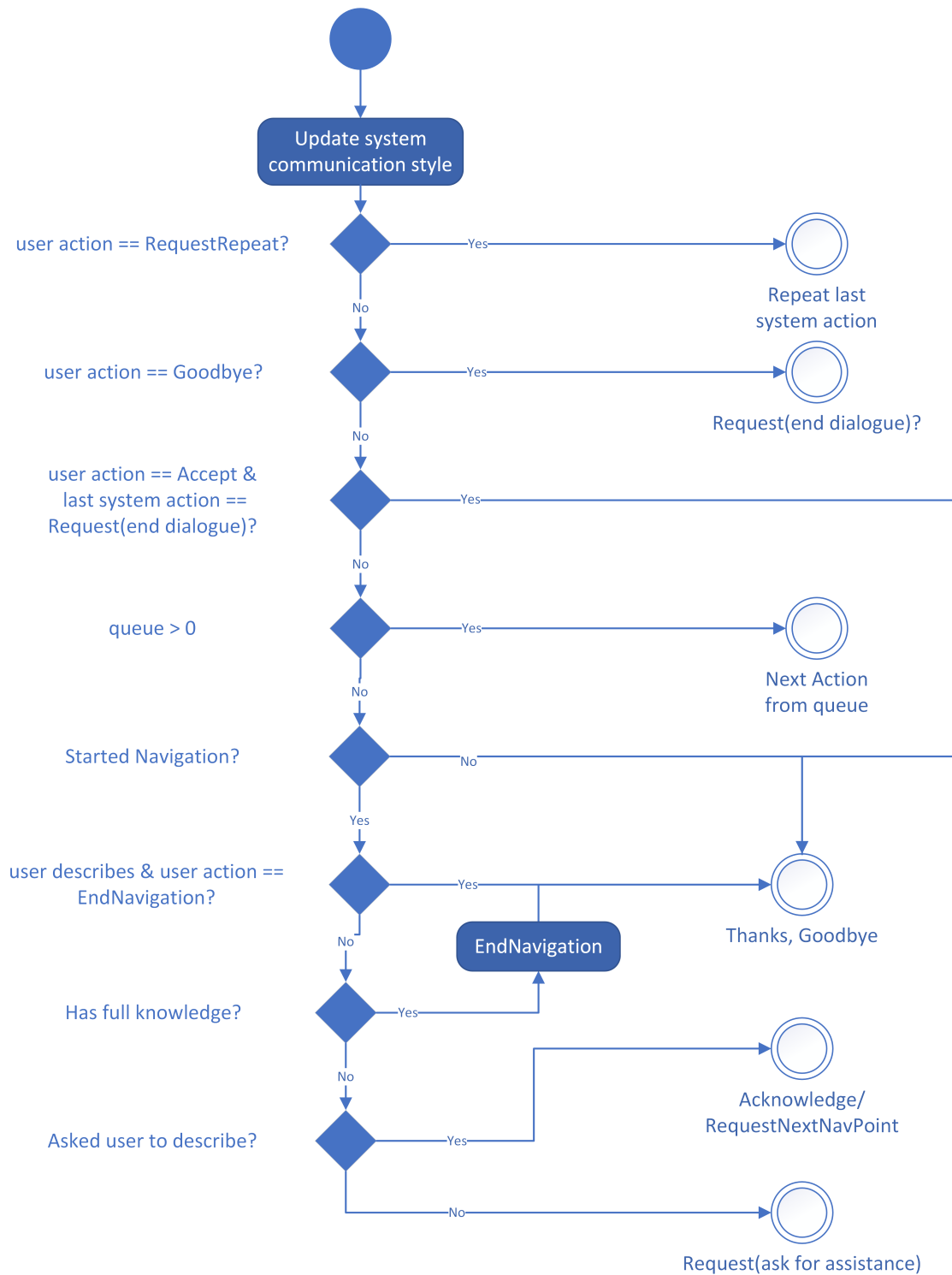


Figure 5.4.: A graphical representation of the dialogue manager decision process.

6. Evaluation

In the following we explain the setting and execution of the study. The results are shown in section 6.3.

Aim of the study The aim of this study is to investigate whether the implemented adaptive communication style strategy for an indoor navigation application positively influences user satisfaction. Furthermore, possible variances between the different knowledge levels of the system will be examined.

6.1. Study design

The study procedure is as follows:

1. Presentation and introduction to the study:

The participant is welcomed and introduced to the study context. A declaration of consent is handed out and the course of the study is explained.

2. The two tasks:

Each task is explained immediately before its start. The participant solves the tasks independently, assistance is limited to the bare essentials. The tasks include:

- a) Small talk dialogue: The subject conducts a small talk dialogue in which the system asks various general questions about the person. The aim is to determine the communication style of the user.
- b) Indoor-Navigation dialogue: The subject sends a navigation request to the system. The system now guides the subject through the university depending on his state of knowledge. The system can have two different knowledge levels, either it knows the complete route and thus navigates the user till the end, or it knows only a part and asks the user for help to end the route. The knowledge level is selected randomly at the beginning of the study and the participant is informed in advance before the task starts. After completing the navigation task, the participant has to fill out a questionnaire on user satisfaction on a tablet. The navigation task is carried out using three different routes in total, with the system using a different communication style (random, fixed or adaptive) for each route.

Within the questionnaire the participants have to rate 47 statements about the navigation dialogue on a five-point Likert scale. We used the SASSI questionnaire[5], which was developed to measure the subjective assessment

6. Evaluation

of speech system interfaces. It includes the following categories: System Response Accuracy, Likeability, Cognitive Demand, Annoyance, Habitability and Speed. In addition, we added further statements on the perception of the system behavior and the system communication style (see Table 6.3).

3. Final questionnaire:

A self-disclosure regarding the participant's own usage of dialogue systems, the preferred dialogue scenario and some free questions have to be completed by the participant. The free questions include:

- Have you noticed any differences between the dialogues? If so, which ones?
- Do you have any additional general remarks about the small talk dialogue?
- Do you have any additional general remarks about the navigation dialogues?

6.2. Execution

6.2.1. Required resources

A tablet and a headset are required as essential resources. The tablet offers the possibility of free movement for the navigation and makes it easy to answer the questionnaires. To ensure a smooth speech dialogue, the headset should on the one hand provide a clear input signal, and on the other hand shield the test person from annoying noises.

6.2.2. Acquisition of volunteers

Several methods were used to acquire volunteers. On the one hand, we asked acquaintances, on the other hand an advertisement was created on the Internet. In addition, the email distributor of the University of Ulm was used.

6.2.3. Pilot study

Before carrying out the actual study, a test person was invited to test the course of the study. This included walking all three ways with the help of the system and filling out the questionnaires. Both knowledge levels of the system and all three communication styles were distributed among the paths. In addition to the general functionality of the system in the right environment, the wifi stability had to be examined. The test run took place on 8 March 2019 and was successfully overall lasting 55 minutes.

6.2.4. Adjustments

The pilot study revealed that, in order to keep the ambient noise in control, the mute button on the headset had to be used when operating the system. The user interface has been adapted to make better use of the speech recognition. At the beginning of the main study, a small talk question was removed because the answers to this question regularly

	# Participants
male	16
female	14
every day	0
several times a week	3
several times a month	7
never	20
20 - 29 years	28
30 - 39 years	2

Table 6.1.: The distribution of study participants regarding their gender, their frequency of use speech based assistants as well as their age.

Knowledge	Dialogue 1	Dialogue 2	Dialogue 3	# Participants
Full	Fixed	Random	Adaptive	5
Half	Fixed	Random	Adaptive	5
Full	Random	Adaptive	Fixed	5
Half	Random	Adaptive	Fixed	5
Full	Adaptive	Fixed	Random	4
Half	Adaptive	Fixed	Random	3
Full	Adaptive	Random	Fixed	1
Half	Adaptive	Random	Fixed	2

Table 6.2.: The distribution of participants among the combinations of dialogue communication styles and knowledge level with the options full knowledge (Full) and half knowledge (Half).

caused speech recognition problems. Moreover, a bug was fixed that caused the system to crash when there was no wifi signal.

6.2.5. Main study

A total of 30 people participated in the study between 11 and 20 March 2019. Table 6.1 shows the distribution of participants in terms of gender, frequency of use speech based assistants and age. A total of six groups were planned to randomize the parameters. There were 3 groups for each of the two states of knowledge. Furthermore, the groups differ in the communication style for the three routes. The complete distribution is shown in Table 6.2. Due to negligence, eight groups were created at the end. However, this does not pose a problem, as the participants have used all three communication styles and the division primarily serves for randomization.

6.3. Results

Table 6.3 shows the results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(f)$ is the mean of the ratings of all dialogues with the fixed communication style strategy, $M(r)$ is the mean of the ratings of all dialogues with the random communication style strategy and $M(a)$ is the mean of the ratings of all dialogues with the adaptive communication style strategy. $D(fa) = |M(f) - M(a)|$ is the difference between $M(f)$ and $M(a)$, and $\alpha(fa)$ is the statistical significance level using the Mann-Whitney-U-Test. The same applies to $D(ra)$ and $\alpha(ra)$.

The statements were classified into the following categories: System Response Accuracy (SRA), Likeability (L), Cognitive Demand (CD), Annoyance (A), Habitability (H), Speed (S), System Behavior (SB) and System Communication Style (SCS).

Unfortunately, there were no significant differences ($\alpha < 0.05$) between the communication style strategies. Nevertheless, some tendencies can be detected. Comparing the fixed and adaptive communication style strategies, it seems that the adaptive system has been perceived as more stable. Overall, it made fewer mistakes, was less repetitive and not as inflexible as the fixed strategy. Moreover, there is a tendency to lose track less when using the adaptive strategy. Nevertheless, there are also contradictions, it was stated that while the adaptive strategy provided more information than the fixed strategy, the users wanted more details from the adaptive system. There is another disagreement regarding the systems speed. The interaction speed between both strategies was assessed similarly, however, the fixed strategy tends to be slower in responding. There are fewer differences between adaptive and random communication style strategies. The interaction with the system with the random strategy tends to be more repetitive and it is harder to not lose track of where you are within the adaptive dialogue. Interestingly, the two strategies differ slightly in terms of the statement "I always knew what the system wanted to tell me.", with the random strategy performing better.

The tendencies do not help us to notice a real difference between the communication style strategies. Even a look at the grouped mean values of the participants ratings in Table 6.4 shows no major differences. However, we can make general conclusions from this table. The average of the SASSI questionnaire statements with a value of about 3.6 shows that our system is not sufficient to satisfy the user. The Speed category is particularly conspicuous, as the fixed communication style strategy achieved the lowest value of 3.2, which can be explained, since the response time of the system varied greatly depending on the WLAN and the classification of the Support Vector Machine. The weaker values of the other categories such as System Response Accuracy, Annoyance and Habitability can be stated by the multiple speech recognition problems the user had. Nevertheless, the acceptance for the system is very high, which is also reflected in the Likeability category. With the exception of two statements, all statements in the category were rated positively. The statements "I would use this system" and "I was able to fix errors easily" are somewhat lower rated, which could have been caused by the speech

6.3. Results

Cat.	Statement	M	σ	$M(f)$	$M(r)$	$M(a)$	$D(fa)$	$\alpha(fa)$	$D(ra)$	$\alpha(ra)$
SRA	The system is accurate	3.67	1.15	3.63	3.73	3.63	0.00	0.9389	0.10	0.8110
	The system is unreliable	2.16	1.14	1.97	2.30	2.20	0.23	0.3389	0.10	0.6391
	The interaction with the system is unpredictable	1.92	1.01	2.07	1.83	1.87	0.20	0.5608	0.03	0.8541
	The system didn't always do what I wanted	2.64	1.18	2.57	2.63	2.73	0.17	0.5725	0.10	0.8484
	The system didn't always do what I expected	2.53	1.14	2.67	2.53	2.40	0.27	0.4144	0.13	0.6377
	The system is dependable.	3.74	1.07	3.77	3.67	3.80	0.03	0.9685	0.13	0.6723
	The system makes few errors.	3.57	1.26	3.30	3.67	3.73	0.43	0.2490	0.07	0.7014
	The interaction with the system is consistent.	4.08	0.89	4.13	4.03	4.07	0.07	0.9103	0.03	0.6766
	The interaction with the system is efficient.	3.47	1.23	3.40	3.40	3.60	0.20	0.4718	0.20	0.3730
L	The system is useful.	4.03	1.10	3.93	4.07	4.10	0.17	0.5047	0.03	0.8990
	The system is pleasant.	3.93	1.12	3.87	3.97	3.97	0.10	0.5517	0.00	0.7360
	The system is friendly.	4.10	1.08	4.00	4.20	4.10	0.10	0.7453	0.10	0.8670
	I enjoyed using the system.	4.06	1.25	4.00	4.13	4.03	0.03	0.9611	0.10	0.8910
	It is clear how to speak to the system.	3.96	0.91	3.93	3.93	4.00	0.07	0.7792	0.07	0.7880
	It is easy to learn to use the system.	4.47	0.80	4.53	4.43	4.43	0.10	0.6199	0.00	0.8860
	I would use this system.	3.48	1.42	3.43	3.53	3.47	0.03	0.9091	0.07	0.7720
	I felt in control of the interaction with the system.	4.01	0.97	4.03	4.03	3.97	0.07	0.6742	0.07	0.6700
CD	I was able to fix errors easily.	3.63	1.05	3.57	3.70	3.63	0.07	0.6728	0.07	0.9690
	I felt confident using the system.	4.00	1.13	4.10	3.93	3.97	0.13	0.6338	0.03	0.7410
	I felt tense using the system.	1.94	1.06	1.93	1.87	2.03	0.10	0.7890	0.17	0.6190
	I felt calm using the system.	4.02	1.19	4.07	4.07	3.93	0.13	0.6363	0.13	0.6530
	A high level of concentration is required when using the system	2.40	1.08	2.47	2.30	2.43	0.03	0.8593	0.13	0.6620
A	The system is easy to use.	4.21	0.98	4.23	4.27	4.13	0.10	0.6990	0.13	0.7000
	The interaction with the system is repetitive.	3.21	1.29	3.37	3.27	3.00	0.37	0.2722	0.27	0.4390
	The interaction with the system is boring.	2.12	1.16	2.17	2.13	2.07	0.10	0.7045	0.07	0.9010
	The interaction with the system is irritating.	2.21	1.23	2.37	2.10	2.17	0.20	0.5025	0.07	0.9130
	The interaction with the system is frustrating.	1.93	1.16	2.13	1.77	1.90	0.23	0.5110	0.13	0.8980
H	The system is too inflexible.	2.64	1.15	2.40	2.77	2.77	0.37	0.2656	0.00	0.9510
	I sometimes wondered if I was using the right word.	3.24	1.24	3.13	3.37	3.23	0.10	0.7710	0.13	0.7050
	I always knew what to say to the system.	3.62	1.06	3.67	3.63	3.57	0.10	0.7798	0.07	0.9380
	I was not always sure what the system was doing.	2.72	1.29	2.77	2.70	2.70	0.07	0.8321	0.00	0.9940
	It is easy to lose track of where you are in an interaction with the system.	2.00	1.06	2.07	2.13	1.80	0.27	0.3276	0.33	0.3070
SD	The interaction with the system is fast.	3.33	1.18	3.23	3.50	3.27	0.03	0.8850	0.23	0.4280
	The system responded too slowly.	2.66	1.16	2.83	2.63	2.50	0.33	0.2394	0.13	0.6580
SB	The system seemed polite.	4.29	1.01	4.17	4.40	4.30	0.13	0.4860	0.10	0.8930
	The system seemed courteous.	3.96	1.13	3.97	3.97	3.93	0.03	0.7967	0.03	0.7970
	The system seemed friendly.	4.18	1.08	4.10	4.23	4.20	0.10	0.4854	0.03	0.7250
	The system seemed professional in its speaking style.	4.08	1.10	4.03	4.03	4.17	0.13	0.5299	0.13	0.6480
SCS	The messages repeated themselves.	2.72	1.27	2.87	2.77	2.53	0.33	0.3765	0.23	0.4030
	The system gave me more details than I needed.	1.88	0.97	1.73	1.93	1.97	0.23	0.2939	0.03	0.9490
	The system spoke too much.	1.80	1.01	1.63	1.93	1.83	0.20	0.5503	0.10	0.5510
	I felt that I had to wait too long for the system to stop talking in order to answer.	1.89	1.06	1.63	2.13	1.90	0.27	0.4829	0.23	0.1680
	The system provided the right amount of information.	3.66	1.18	3.73	3.70	3.53	0.20	0.5112	0.17	0.4910
	I got the information I wanted.	3.98	1.10	4.03	3.83	4.07	0.03	0.9811	0.23	0.3260
	I would have liked more details.	2.50	1.29	2.30	2.43	2.77	0.47	0.1567	0.33	0.3150
	I always knew what the system wanted to tell me.	3.93	1.19	3.87	4.17	3.77	0.10	0.8886	0.40	0.3750
	The system didn't answer my questions accurately.	2.13	1.11	2.37	1.93	2.10	0.27	0.3483	0.17	0.7310

Table 6.3.: The results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(f)$ is the mean of the ratings of all dialogues with the fixed communication style strategy, $M(r)$ is the mean of the ratings of all dialogues with the random communication style strategy and $M(a)$ is the mean of the ratings of all dialogues with the adaptive communication style strategy. $D(fa) = |M(f) - M(a)|$ is the difference between $M(f)$ and $M(a)$, and $\alpha(fa)$ is the statistical significance level using the Mann-Whitney-U-Test. The same applies to $D(ra)$ and $\alpha(ra)$.

6. Evaluation

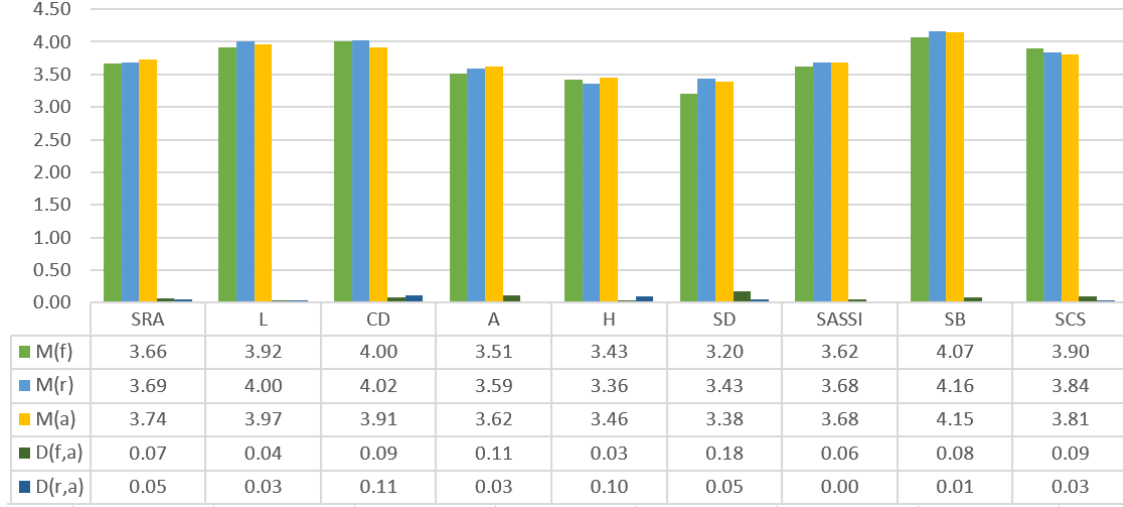


Table 6.4.: The results of the participants’ ratings grouped by the different categories of the questionnaire, whereby the rating scale has been inverted for the negatively formulated statements so that the optimal rating is five for every statement. $M(f)$ is the mean of the ratings of all dialogues with the fixed communication style strategy, $M(r)$ is the mean of the ratings of all dialogues with the random communication style strategy and $M(a)$ is the mean of the ratings of all dialogues with the adaptive communication style strategy.

recognition problems, the general limitation of the system and the usage of a speech interface. One participant mentioned that he likes the overall system, but he would not use it with a speech interface, because he feels uncomfortable with it. Nevertheless, the category Cognitive Demand has positive characteristics, which indicates a simple handling of the system and the System Behavior reached the highest mean values of the categories. According to this, the system is perceived as very sociable. The statements about the system communication style also reached a positive value, but in the overall category there are no major differences between the communication style strategies observed.

We already mentioned some causes for the insufficient results. Moreover, we suspect that the small number of respondents and the large number of parameters have a huge impact on the results. When filling in the question concerning the preference of a navigation dialogue, it was observed that it was very difficult for the participants to make a decision. They often asked how to rate the best one. Moreover, one participant even said within the general remarks: ”I found the dialogues all quite good and it was hard for me to rate one as preferred”. The distribution of the preferred dialogue can be seen in Table 6.5. Table 6.5a shows the selection of the preferred dialogue run and it can be seen that the third dialogue was chosen more often, which can be explained due to the fact, that the users got used to the system and thus had less problems in interacting with it. Also when looking at Table 6.5b we see an equal distribution over the communication style strategies. This suggests that the participants were generally more involved in the

	# Participants		# Participants
First Dialogue	9	Fixed	10
Second Dialogue	7	Random	10
Third Dialogue	14	Adaptive	10

(a) The preferences of the dialogues based on their order within the study.

(b) The preferences mapped to the communication style strategies based on the participants' group.

Table 6.5.: The distribution of the communication style strategies of the preferred dialogues showing how many study participants preferred each option.

navigation task and the basic interaction with the system than in noticing differences within the communication style strategies. Since there were no particular findings, the mean values of the preferred and not preferred dialogues can be found in the Appendix in Table A.2.

In addition, there were many disruptive factors such as the inconsistent WLAN connection, which varied depending on the time. Moreover, some test persons did not have any problems with the WLAN and others had difficulties because they had to pause multiple times within the dialogues. Besides, the speech recognition was defective. This is a general problem when evaluating a dialogue management strategy, as you can only test the whole system at once. In the case of our system, the natural language understanding was trained without sufficient data, resulting in a small variety and inconsistency of accepted spoken input patterns. For a system that aims at natural communication this is of course an inconvenience, since the user adapts to the characteristics of the system. These factors promote a negative attitude towards the systems accuracy and dependability.

As already mentioned, another reason for the results might be the many variables used in comparison to the amount of participants. We assume that with a higher number of participants, better results can be achieved and suspect that the other factors, like knowledge and routes, play a more prominent role than the communication style strategies. Therefore, we will now take a closer look to the different states of knowledge to see if there were any significant differences.

Table 6.6 shows the results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(k_f)$ is the mean of the ratings of all dialogues with the system having full knowledge and $M(k_p)$ is the mean of the ratings of all dialogues with the system having partial knowledge. $D = |M(k_f) - M(k_p)|$ is the difference between $M(k_f)$ and $M(k_p)$, and α is the statistical significance level using the Mann-Whitney-U-Test. The differences that are statistically significant ($\alpha < 0.05$) are marked (*).

6. Evaluation

Cat.	Statement	M	σ	$M(k_f)$	$M(k_p)$	D	α
SRA	The system is accurate	3.67	1.15	3.71	3.62	0.09	0.7750
	The system is unreliable	2.16	1.14	2.24	2.07	0.17	0.3280
	The interaction with the system is unpredictable	1.92	1.01	1.98	1.87	0.11	0.4740
	The system didn't always do what I wanted	2.64	1.18	2.69	2.60	0.09	0.7000
	The system didn't always do what I expected	2.53	1.14	2.58	2.49	0.09	0.4630
	The system is dependable.	3.74	1.07	3.76	3.73	0.03	0.7320
	The system makes few errors.	3.57	1.26	3.40	3.73	0.33	0.1340
	The interaction with the system is consistent.	4.08	0.89	3.98	4.18	0.20	0.1100
L	The interaction with the system is efficient.	3.47	1.23	3.44	3.49	0.05	0.6690
	The system is useful.	4.03	1.10	3.89	4.18	0.29	0.0720
	The system is pleasant.	3.93	1.12	3.80	4.07	0.27	0.3880
	The system is friendly.	4.10	1.08	4.16	4.04	0.12	0.9720
	I enjoyed using the system.	4.06	1.25	3.82	4.29	0.47	0.0830
	It is clear how to speak to the system.	3.96	0.91	4.13	3.78	0.35	0.0840
	It is easy to learn to use the system.	4.47	0.80	4.51	4.42	0.09	0.9040
	I would use this system.	3.48	1.42	3.22	3.73	0.51	0.0610
CD	I felt in control of the interaction with the system.	4.01	0.97	3.87	4.16	0.29	0.0950
	I was able to fix errors easily.	3.63	1.05	3.49	3.78	0.29	0.0840
	I felt confident using the system.	4.00	1.13	4.02	3.98	0.04	0.9040
	I felt tense using the system.	1.94	1.06	2.02	1.87	0.15	0.4490
	I felt calm using the system.	4.02	1.19	4.04	4.00	0.04	0.9140
A	A high level of concentration is required when using the system	2.40	1.08	2.49	2.31	0.18	0.4900
	The system is easy to use.	4.21	0.98	4.22	4.20	0.02	0.9890
	The interaction with the system is repetitive.	3.21	1.29	3.40	3.02	0.38	0.1480
	The interaction with the system is boring.	2.12	1.16	2.22	2.02	0.20	0.3770
	The interaction with the system is irritating.	2.21	1.23	2.40	2.02	0.38	0.1190
H	The interaction with the system is frustrating.	1.93	1.16	2.09	1.78	0.31	0.1480
	The system is too inflexible.	2.64	1.15	2.93	2.36	0.57	0.0100*
	I sometimes wondered if I was using the right word.	3.24	1.24	3.02	3.47	0.45	0.0960
	I always knew what to say to the system.	3.62	1.06	3.73	3.51	0.22	0.3630
SD	I was not always sure what the system was doing.	2.72	1.29	2.71	2.73	0.02	0.9900
	It is easy to lose track of where you are in an interaction with the system.	2.00	1.06	2.16	1.84	0.32	0.1650
	The interaction with the system is fast.	3.33	1.18	3.24	3.42	0.18	0.3880
SB	The system responded too slowly.	2.66	1.16	2.67	2.64	0.03	0.8740
	The system seemed polite.	4.29	1.01	4.24	4.33	0.09	0.5100
	The system seemed courteous.	3.96	1.13	3.89	4.02	0.13	0.2500
	The system seemed friendly.	4.18	1.08	4.16	4.20	0.04	0.6080
SCS	The system seemed professional in its speaking style.	4.08	1.10	3.93	4.22	0.29	0.1980
	The messages repeated themselves.	2.72	1.27	2.56	2.89	0.33	0.2230
	The system gave me more details than I needed.	1.88	0.97	2.11	1.64	0.47	0.003*
	The system spoke too much.	1.80	1.01	1.93	1.67	0.26	0.1250
	I felt that I had to wait too long for the system to stop talking in order to answer.	1.89	1.06	1.89	1.89	0.0	0.7380
	The system provided the right amount of information.	3.66	1.18	3.62	3.69	0.07	0.6680
	I got the information I wanted.	3.98	1.10	3.84	4.11	0.27	0.042*
	I would have liked more details.	2.50	1.29	2.44	2.56	0.12	0.6800
	I always knew what the system wanted to tell me.	3.93	1.19	3.71	4.16	0.45	0.0760
	The system didn't answer my questions accurately.	2.13	1.11	2.18	2.09	0.09	0.3990

Table 6.6.: The results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(k_f)$ is the mean of the ratings of all dialogues with the system having full knowledge and $M(k_p)$ is the mean of the ratings of all dialogues with the system having partial knowledge. $D = |M(k_f) - M(k_p)|$ is the difference between $M(k_f)$ and $M(k_p)$, and α is the statistical significance level using the Mann-Whitney-U-Test. The differences that are statistically significant ($\alpha < 0.05$) are marked (*).

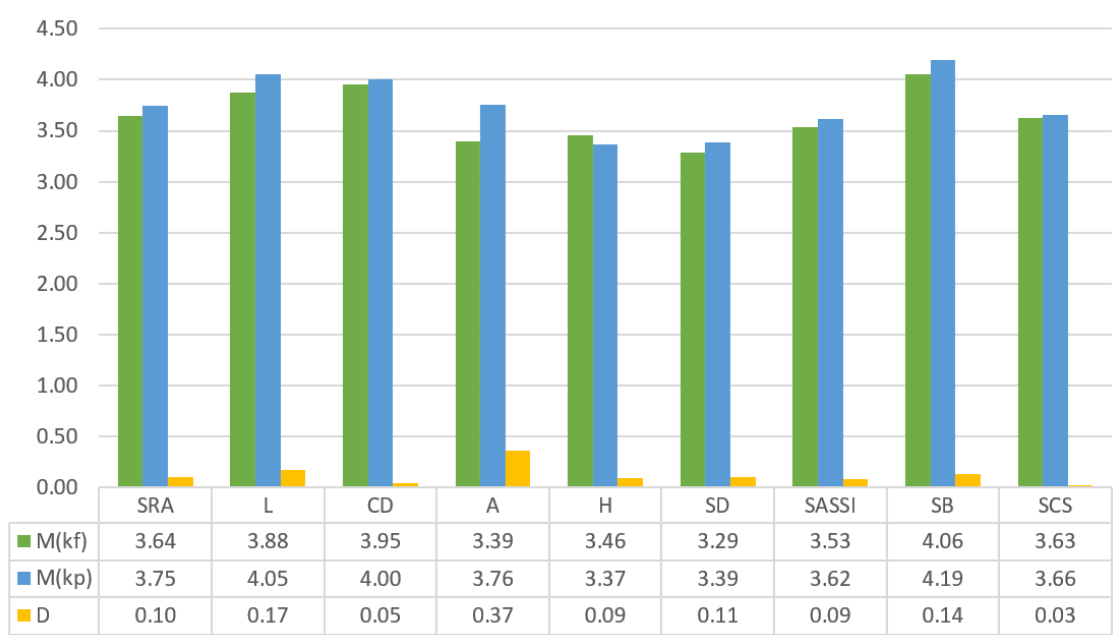


Table 6.7.: The results of the participants' ratings grouped by the different categories of the questionnaire, whereby the rating scale has been inverted for the negatively formulated statements so that the optimal rating is five for every statement. $M(k_f)$ is the mean of the ratings of all dialogues with full knowledge and $M(k_p)$ is the mean of the ratings of all dialogues with partial knowledge, $D = M(k_f) - M(k_p)$

In fact, the two different states of knowledge of the system provide significant differences. Interestingly, the system with partial knowledge is seen as more flexible than the system with full knowledge. The two other significant differences in the statements regarding the level of detail and the information received are plausible, since the system with partial knowledge can provide less information than the system with full knowledge. Apart from the significant differences, there are also some visible tendencies here. Especially in the category Likeability a clear difference can be noticed. The system with the half-knowledge receives higher ratings in the corresponding statements. Also in the average of the category Likeability (see Table 6.7) a clear tendency of 0.3 can be recognized. Another interesting aspect are the statements "The interaction with the system is repetitive" and "The messages repeated themselves". In the first statement, the system with the half-knowledge performs better and in the second the system with the full knowledge performs better. This means that the system with the half-knowledge probably repeats itself more in its statements (especially in the part where the user has to navigate) and yet the interaction is seen as not as repetitive as with the system with the full knowledge. A further point is visible in the statements of the Habitability category. Here, users often seem to be looking for formulations when they have to navigate, which makes

6. Evaluation

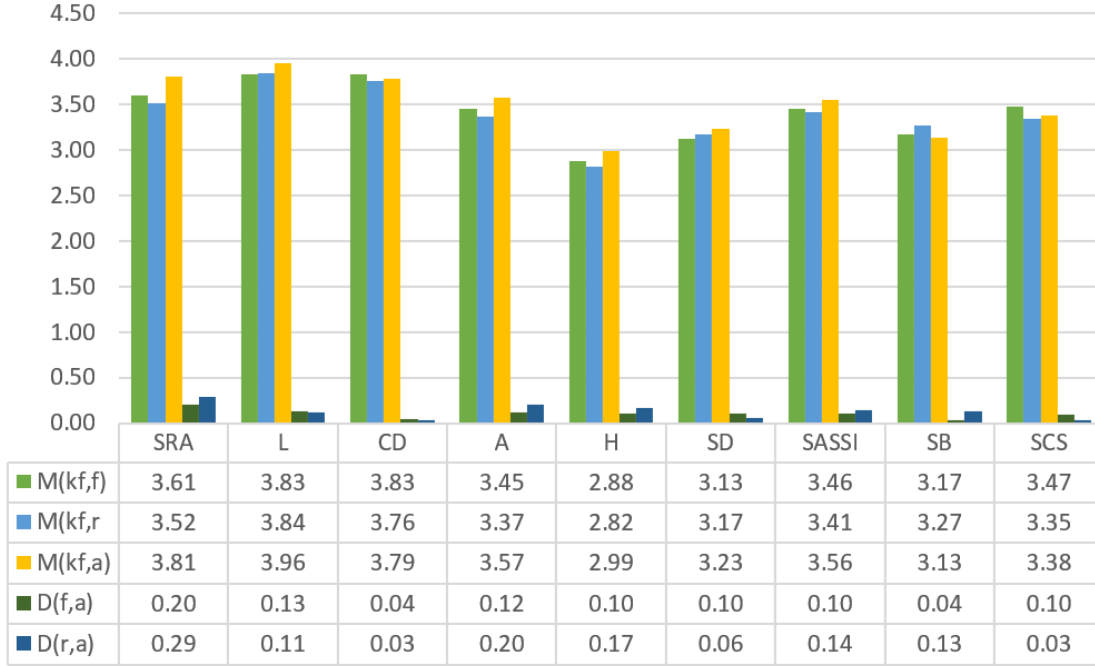


Table 6.8.: The results of the participants' ratings grouped by the different categories of the questionnaire, whereby the rating scale has been inverted for the negatively formulated statements so that the optimal rating is five for every statement. $M(k_f, x)$ is the mean of the ratings of all dialogues with full knowledge and the communication style strategy x ($f = fixed, r = random, a = adaptive$), $D(fa) = M(k_f, f) - M(k_f, a)$ and $D(ra) = M(k_f, r) - M(k_f, a)$ are the differences between the mean values.

sense. Furthermore, the dialogue flow with the system with half knowledge seems to be clearer than with the system with full knowledge. This corresponds with the statement from the category System Communication Style: "I always knew what the system wanted to tell me", whereby the system with the half-knowledge also obtained a better result. This confirms our basic idea that two levels of knowledge of a system lead to different interactions with the user so that the user behaves and feels differently.

Since we have found significantly difference between the two states of knowledge, our next step is to examine whether the individual communication style strategies show significant differences in relation to one state of knowledge. First we look at the communication styles in relation to the full knowledge. Since we don't want to go into all the tendencies of each individual statement here, we will only show the significant results and the chart of the distribution across the individual categories (see Table 6.8). The related Table for all mean values of the statements A.3 can be found in the appendix.

The only significant result was the statement "The system didn't answer my questions ac-

curately”, in which the adaptive communication style strategy was rated more positively than the fixed or random strategy. Table 6.8 shows that the adaptive communication style strategy performs better in some aspects. In relation to the SASSI questionnaire, the adaptive style leads with a slight tendency and generally the dialogue seems to be more accurate. With regard to the communication style, it is interesting to see that the fixed style was more popular. However, this can also be influenced by the directness. Since the classification was not sufficient, some direct dialogues quickly became indirect. The indirect style often caused confusion among the participants and the system did not react to their questions due to its restriction.

Table 6.9 shows the mean values of the categories concerning the communication style strategies of the system with partial knowledge. In contrast to the values of the system with full knowledge, the adaptive strategy of the semi-knowing system, apart from the category System Behavior, tends to perform less satisfactorily. The random strategy is among the best in this category. This could be explained by the fact that the random strategy brings more variety in the systems statements, while the user describes the path.

Now we take a look at the answers to the open questions of the general questionnaire. When asked whether differences between dialogues were detected, the participants recognized differences in the level of detail of the system statements. This suggests that the differences between concise and elaborate were perceived. A few participants also noticed the difference in directness.

In general, the small talk dialogue was not positively perceived, as it is very restrictive. The participants wanted a more realistic interaction with the system, with the system responding to the spoken statements of the users. One participant described the dialogue as generally not very realistic. He said: ”The system should not try to act like a human being, but should be aware of its mechanical nature.”

A lot of the comments on the navigation dialogues refer to improve the speech recognition. In addition, it was noted that the system could give the next instruction on its own when the user has completed an intermediate step. Another participant said that the system had spoken too much between the user’s descriptions and a simple confirmation would have been enough. Furthermore, alternative descriptions for the individual steps and a ”back” function were described as helpful. The system should react to position information from the user, e.g. the user says ”I have arrived at the end of the corridor”, the system should now initiate the next step on its own.

6. Evaluation

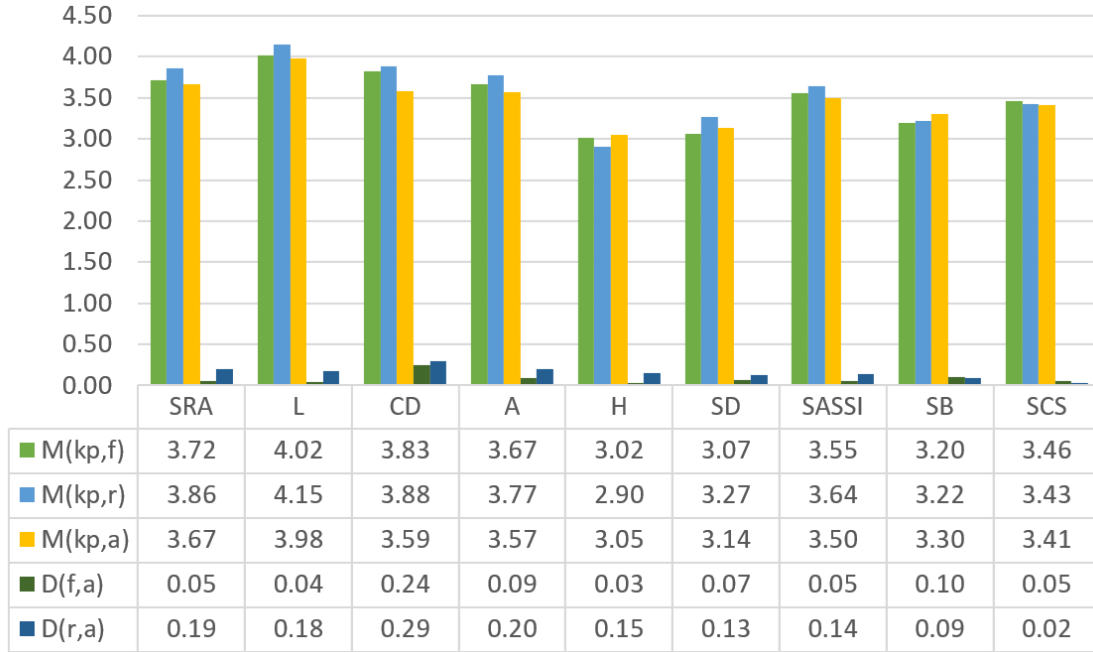


Table 6.9.: The results of the participants' ratings grouped by the different categories of the questionnaire, whereby the rating scale has been inverted for the negatively formulated statements so that the optimal rating is five for every statement. $M(k_p, x)$ is the mean of the ratings of all dialogues with partial knowledge and the communication style strategy x ($f = fixed, r = random, a = adaptive$), $D(fa) = M(k_p, f) - M(k_p, a)$ and $D(ra) = M(k_p, r) - M(k_p, a)$ are the differences between the mean values.

7. Conclusion

Within this work we have seen that the communication styles elaborateness and directness have an influence on the user's satisfaction in spoken dialogue systems. We have examined three different kinds of communication style strategies including a fixed, a random and an adaptive strategy. Moreover, we developed a formula for the adaptive strategy, to adapt the system's communication style to the user's communication style. We developed an indoor navigation spoken dialogue system, which has these three different communication style strategies and two states of knowledge implemented. We decided to use two different kinds of knowledge, because we wanted to investigate the different styles of interaction with the system. The system will navigate the user based on scripts through the University of Ulm and can be expanded with more scripts easily.

For the scripting a data collection study was carried out in which 97 participants described the selected routes with the help of videos, which were recorded beforehand. To conduct the study, a web application was implemented that enabled the participants to create a list of descriptions. Furthermore, the results of the study were processed and three scripts with four different communication styles were created on the basis of the collected descriptions. Finally, the system was evaluated in a user study with 30 participants.

In summary, one can say that the system has basically made a good performance (SASSI mean value of 3.6 or more). The application of an indoor navigation was generally seen as very helpful. The deliberate limitations of the system explain many negative aspects, which can be very helpful for further development. Above all, speech recognition needs to be improved.

Unfortunately, no significant differences could be found between the communication style strategies, which can be explained by the many variables of the study. For a further study, either more participants should be included or some parameters should be simplified.

For example, one could use only one route and let the participants walk along this route with the system using different communication style strategies. Furthermore, the level of knowledge should be reduced to one and the question arises, whether the application of indoor navigation is a suitable application to investigate the differences in communication style strategies. In addition, it became visible during the study that an adaptive approach, which adapts to the style of the user, is not necessarily effective. Since, especially in the case of the indoor navigation task, the users often want to get to their information quickly, and therefore may be curt, but they do not necessarily want a short answer from the system.

In the future, the system should manage the navigation information as semantic graphs.

7. Conclusion

This allows the system to answer questions of the user and to generate the communication style based on semantic representation. Lucena et al.[10] already showed how a possible realization of this could look like. Within their work they tried to generate natural language that has a high factor of humanlikeness by use of semantic representation.

With regard to the communication styles elaborateness and directness it also makes sense to further optimize the classification of these. The directness in particular requires further research into which factors really influence it and how it can be better classified. Furthermore, one can consider further gradations of elaborateness in order to offer the user even more variety.

The dialogue management should also be adapted. It would make sense to try out a new strategy such as reinforcement learning. However, dialogue data is needed to train this meaningfully. Furthermore, the dialogue actions should become more detailed in general and have more semantic content, e.g. entities, to enable a better dialogue flow.

Overall, within this work we were able to demonstrate the potential of an indoor navigation application with voice interface and to highlight some problem areas, which could be solved in the future.

Bibliography

- [1] Robert Belvin, Ron Burns, and Cheryl Hein. Development of the hrl route navigation dialogue system. In *Proceedings of the first international conference on Human language technology research*, pages 1–5. Association for Computational Linguistics, 2001.
- [2] Markus Berg. Survey on spoken dialogue systems: User expectations regarding style and usability. In *14th International PhD Workshop OWD*, 2012.
- [3] Susan E Brennan and Justina O Ohaeri. Effects of message style on users’ attributions toward agents. In *Conference on Human Factors in Computing Systems: Conference companion on Human factors in computing systems*, volume 24, pages 281–282, 1994.
- [4] Kate Forbes-Riley and Diane J Litman. Modelling user satisfaction and student learning in a spoken dialogue tutoring system with generic, tutoring, and user affect parameters. In *Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, pages 264–271. Association for Computational Linguistics, 2006.
- [5] Kate S Hone and Robert Graham. Towards a tool for the subjective assessment of speech system interfaces (sassi). *Natural Language Engineering*, 6(3-4):287–303, 2000.
- [6] Topi Hurtig. A mobile multimodal dialogue system for public transportation navigation evaluated. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, pages 251–254. ACM, 2006.
- [7] Ivana Kruijff-Korbayová, Ciprian Gerstenberger Olga Kukina, and Jan Schehl. Generation of output style variation in the sammie dialogue system. In *Proceedings of the Fifth International Natural Language Generation Conference*, pages 129–137. Association for Computational Linguistics, 2008.
- [8] Ivana Kruijff-Korbayová and Olga Kukina. The effect of dialogue system output style variation on users’ evaluation judgments and input style. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*, pages 190–197. Association for Computational Linguistics, 2008.
- [9] Diane J Litman and Shimei Pan. Designing and evaluating an adaptive spoken dialogue system. *User Modeling and User-Adapted Interaction*, 12(2-3):111–137, 2002.

Bibliography

- [10] Diego Jesus de Lucena, Daniel Bastos Pereira, and Ivandré Paraboni. From semantic properties to surface text: The generation of domain object descriptions. *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*, 14(45), 2010.
- [11] François Mairesse and Marilyn A Walker. Towards personality-based user adaptation: psychologically informed stylistic language generation. *User Modeling and User-Adapted Interaction*, 20(3):227–278, 2010.
- [12] Juliana Miehle, Wolfgang Minker, and Stefan Ultes. Exploring the impact of elaborateness and indirectness on user satisfaction in a spoken dialogue system. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization (UMAP)*, pages 165–172. ACM, July 2018.
- [13] Juliana Miehle, Wolfgang Minker, and Stefan Ultes. What causes the differences in communication styles? a multicultural study on directness and elaborateness. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*, 2018.
- [14] Juliana Miehle, Koichiro Yoshino, Louisa Pragst, Stefan Ultes, Satoshi Nakamura, and Wolfgang Minker. Cultural communication idiosyncrasies in human-computer interaction. In *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 74–79, 2016.
- [15] C Nass and S Brave. Should voice interfaces say” i”? recorded and synthetic voice interfaces’ claims to humanity, chapter 10, 2005.
- [16] Lisa Ran, Sumi Helal, and Steve 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*, pages 23–30. IEEE, 2004.
- [17] Michaela Regneri, Alexander Koller, and Manfred Pinkal. Learning script knowledge with web experiments. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 979–988, 2010.
- [18] Kazuki Sakai, Ryuichiro Higashinaka, Yuichiro Yoshikawa, Hiroshi Ishiguro, and Junji Tomita. Introduction method for argumentative dialogue using paired question-answering interchange about personality. In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue*, pages 70–79, 2018.
- [19] Martin Werner and Moritz Kessel. *Organisation of Indoor Navigation data from a data query perspective*. IEEE, 2010.

A. Additional Material

A.1. The Scripts

A.1.1. The Small Talk Script

```
<smalltalk>
  <question-answer-pair>
    <question>
      <concise_direct>How has your day been so far?</concise_direct>
    </question>
    <answer>
      <concise_direct>My day has been good. I had a lot of good
        talks today!</concise_direct>
    </answer>
  </question-answer-pair>
  <question-answer-pair>
    <question>
      <concise_direct>How do you like the weather?</concise_direct>
    </question>
    <answer>
      <concise_direct>I like this kind of weather.</concise_direct>
    </answer>
  </question-answer-pair>
  <question-answer-pair>
    <question>
      <concise_direct>What hobbies do you have?</concise_direct>
    </question>
    <answer>
      <concise_direct>I do a lot of outdoor sports!</concise_direct>
    </answer>
  </question-answer-pair>
  <question-answer-pair>
    <question>
      <concise_direct>Do you have plans for the evening?</
        concise_direct>
      <elaborate>The day is almost over! Do you have any interesting
        plans for the evening?</elaborate>
    </question>
```

A. Additional Material

```
<answer>
  <concise_direct>I guess i will stay at home and relax.</
    concise_direct>
</answer>
</question-answer-pair>
<question-answer-pair>
  <question>
    <concise_direct>Do you like sports?</concise_direct>
  </question>
  <answer>
    <concise_direct>I really like watching sports.</concise_direct
      >
  </answer>
</question-answer-pair>
<question-answer-pair>
  <question>
    <concise_direct>Do you like animals?</concise_direct>
  </question>
  <answer>
    <concise_direct>I love animals.</concise_direct>
  </answer>
</question-answer-pair>
</smalltalk>
```

A.1.2. The Description Script of the first Route

```
<path>
  <pathelement id='1' nextElement='2'>
    <concise_direct>Go straight to the stairs.</concise_direct>
    <elaborate_direct>Go straight and turn left near the stand
      containing magazines in order to reach the stairs.
    </elaborate_direct>
    <concise_indirect>Find the stairs.</concise_indirect>
    <elaborate_indirect>Find the stairs to the left of the stand
      containing magazines</elaborate_indirect>
  </pathelement>
  <pathelement id='2' nextElement='3'>
    <concise_direct>Go down stairs</concise_direct>
    <elaborate_direct>Go down the stairs and pass the sign saying "
      Hochschulsport".</elaborate_direct>
    <concise_indirect>The path continues down the stairs.</
      concise_indirect>
    <elaborate_indirect>The path continues down the stairs, passing a
      sign saying "Hochschulsport".
  </pathelement>
```



```

    </elaborate_indirect>
</pathelement>
<pathelement id='3' nextElement='4'>
    <concise_direct>Enter the door and turn right</concise_direct>
    <elaborate_direct>After reaching the bottom of the stairs, enter
        the door and turn right.</elaborate_direct>
    <concise_indirect>you can see a door.</concise_indirect>
    <elaborate_indirect>After reaching the bottom of the stairs, you
        can see a door</elaborate_indirect>
</pathelement>
<pathelement id='4' nextElement='5'>
    <concise_direct>Go down the ramp.</concise_direct>
    <elaborate_direct>Cross the room until you reach a ramp. Turn left
        and go down the ramp.</elaborate_direct>
    <concise_indirect>There is a ramp.</concise_indirect>
    <elaborate_indirect>There is a ramp at the other end of the room
        .</elaborate_indirect>
</pathelement>
<pathelement id='5' nextElement='6'>
    <concise_direct>Turn right at the first cross</concise_direct>
    <elaborate_direct>Turn right at the first cross after passing the
        doorway.</elaborate_direct>
    <concise_indirect>your next turn is to the right.</
        concise_indirect>
    <elaborate_indirect>after passing the doorway, your next turn is
        to the right.</elaborate_indirect>
</pathelement>
<pathelement id='6' nextElement='7'>
    <concise_direct>Go straight and turn right at the elevator</
        concise_direct>
    <elaborate_direct>Go straight to the end of the corridor and turn
        right when you see an elevator.
    </elaborate_direct>
    <concise_indirect>your way leads you to the right at the end of
        the corridor</concise_indirect>
    <elaborate_indirect>your way leads you to the right at the end of
        the corridor where you can see an elevator
    </elaborate_indirect>
</pathelement>
<pathelement id='7' nextElement='8'>
    <concise_direct>Press the switch to open the door and go through
        the door</concise_direct>
    <elaborate_direct>Press the white switch that says "Rampe A 026

```

A. Additional Material

```
    Zentrallager Laborbedarf" to open the door. Go
    through the blue door which you just opened.
  </elaborate_direct>
  <concise_indirect>The door opens when pressing the switch. The
    path continues through the door.
  </concise_indirect>
  <elaborate_indirect>The door opens when pressing the white switch
    that says "Rampe A 026 Zentrallager
    Laborbedarf". The path continues through the blue door which
    has just been opened.
  </elaborate_indirect>
</pathelement>
<pathelement id='8' nextElement='9'>
  <concise_direct>Turn left</concise_direct>
  <elaborate_direct>Go straight down the hallway and turn left.</
    elaborate_direct>
  <concise_indirect>your next turn is to the left</concise_indirect>
  <elaborate_indirect>your next turn is to the left at the end of
    the corridor</elaborate_indirect>
</pathelement>
<pathelement id='9' nextElement='10'>
  <concise_direct>Continue through the clear plastic door and turn
    right.</concise_direct>
  <elaborate_direct>Continue through the clear plastic door and turn
    right. You will see a set of white doors.
  </elaborate_direct>
  <concise_indirect>your way leads you to the right after the clear
    plastic door</concise_indirect>
  <elaborate_indirect>You will see a set of white doors when turning
    to the right after the clear plastic door
  </elaborate_indirect>
</pathelement>
<pathelement id='10' nextElement='11'>
  <concise_direct>Go through the white door.</concise_direct>
  <elaborate_direct>Go straight and enter the white door at the end
    of the corridor.</elaborate_direct>
  <concise_indirect>There is a white door.</concise_indirect>
  <elaborate_indirect>There is a white door at the end of the
    corridor.</elaborate_indirect>
</pathelement>
<pathelement id='11' nextElement='12'>
  <concise_direct>Turn right</concise_direct>
  <elaborate_direct>Turn right after passing the door.</
```

```

        elaborate_direct>
    <concise_indirect>your next turn is to the right.</
        concise_indirect>
    <elaborate_indirect>your next turn is to the right after passing
        the door.</elaborate_indirect>
</pathelement>
<pathelement id='12' nextElement=''>
    <concise_direct>Pass the lockers and turn left.</concise_direct>
    <elaborate_direct>Slightly turn left and go through another set of
        doors next to a pinboard. Pass the lockers
        and turn left.
    </elaborate_direct>
    <concise_indirect>Your destination is on the left after passing
        the lockers.</concise_indirect>
    <elaborate_indirect>Your destination is on the left after entering
        another set of doors next to a pinboard and
        passing the lockers.
    </elaborate_indirect>
</pathelement>
</path>

```

A.1.3. The Description Script of the second Route

```

<path>
    <pathelement id='1' nextElement='2'>
        <concise_direct>continue down the corridor</concise_direct>
        <elaborate_direct>continue down the corridor, passing the notice
            boards and fire extinguisher</elaborate_direct>
        <concise_indirect>the path continues down the corridor</
            concise_indirect>
        <elaborate_indirect>the path continues down the corridor, passing
            the notice boards and fire extinguisher
        </elaborate_indirect>
    </pathelement>
    <pathelement id='2' nextElement='3'>
        <concise_direct>pass through the double doors and go up the stairs
            </concise_direct>
        <elaborate_direct>Pass through the set of double doors at the end
            of the corridor and go up the stairs.
        </elaborate_direct>
        <concise_indirect>Find the stairs behind the double doors.</
            concise_indirect>
        <elaborate_indirect>Find the stairs behind the double doors at the
            end of the corridor</elaborate_indirect>
    </pathelement>
</path>

```

A. Additional Material

```
</pathelement>
<pathelement id='3' nextElement='4'>
  <concise_direct>Turn right.</concise_direct>
  <elaborate_direct>At the top of the stairs turn right.</
    elaborate_direct>
  <concise_indirect>your next turn is to the right</concise_indirect>
  >
  <elaborate_indirect>your next turn is to the right at the top of
    the stairs.</elaborate_indirect>
</pathelement>
<pathelement id='4' nextElement='5'>
  <concise_direct>go through the double doors</concise_direct>
  <elaborate_direct>go through the double doors with the big windows
    on the other side.</elaborate_direct>
  <concise_indirect>there are double doors</concise_indirect>
  <elaborate_indirect>there are double doors with big windows on the
    other side.</elaborate_indirect>
</pathelement>
<pathelement id='5' nextElement='6'>
  <concise_direct>continue down the corridor and follow it turning
    right.</concise_direct>
  <elaborate_direct>Continue down the corridor with the blue windows
    on the left-hand side. Pass the white lockers
    and turn right at the end of the corridor
  </elaborate_direct>
  <concise_indirect>the path continues down the corridor with the
    next section starting on your right
  </concise_indirect>
  <elaborate_indirect>The path continues down the corridor with the
    blue windows on the left-hand side. The next
    section starts on your right at the end of the corridor after
    passing the white lockers
  </elaborate_indirect>
</pathelement>
<pathelement id='6' nextElement='7'>
  <concise_direct>Turn left.</concise_direct>
  <elaborate_direct>walk forward to the green wall and then turn
    left.</elaborate_direct>
  <concise_indirect>your next turn is to the left</concise_indirect>
  <elaborate_indirect>your next turn is to the left when you see a
    green wall</elaborate_indirect>
</pathelement>
<pathelement id='7' nextElement='8'>
```

```

    <concise_direct>Turn right towards a green wall.</concise_direct>
    <elaborate_direct>Follow the corridor and turn right towards
        another green wall with the number 2 on it.
    </elaborate_direct>
    <concise_indirect>you see a green wall after turning right</
        concise_indirect>
    <elaborate_indirect>you see a green wall with the number 2 on it
        after following the corridor and turning right
    </elaborate_indirect>
</pathelement>
<pathelement id='8' nextElement='9'>
    <concise_direct>Turn right</concise_direct>
    <elaborate_direct>Turn right when you reach the wall</
        elaborate_direct>
    <concise_indirect>the path continues to the right</
        concise_indirect>
    <elaborate_indirect>when reaching the wall, the path continues to
        the right</elaborate_indirect>
</pathelement>
<pathelement id='9' nextElement='10'>
    <concise_direct>go towards the wall with the number 2 on it.</
        concise_direct>
    <elaborate_direct>Continue down the corridor and go towards the
        wall with the giant number 2 on it.
    </elaborate_direct>
    <concise_indirect>find the wall with the number 2 on it.</
        concise_indirect>
    <elaborate_indirect>find the wall with the giant number 2 on it at
        the other end of the corridor
    </elaborate_indirect>
</pathelement>
<pathelement id='10' nextElement='11'>
    <concise_direct>pass through two sets of double doors</
        concise_direct>
    <elaborate_direct>Follow the corridor and pass through the two
        sets of big glass double doors</elaborate_direct>
    <concise_indirect>There are two sets of double doors</
        concise_indirect>
    <elaborate_indirect>there are two sets of big glass double doors
        after following the corridor
    </elaborate_indirect>
</pathelement>
<pathelement id='11' nextElement='12'>

```

A. Additional Material

```
<concise_direct>Turn right and go to the lift</concise_direct>
<elaborate_direct>Turn right at the staircase and go to the lift
  with the green doors.</elaborate_direct>
<concise_indirect>your next turn is to the right. Find the lift.</
  concise_indirect>
<elaborate_indirect>your next turn is to the right when you see
  the staircase. Find the lift with the green
  doors.
</elaborate_indirect>
</pathelement>
<pathelement id='12' nextElement='13'>
  <concise_direct>press the up arrow and go to level 4</
    concise_direct>
  <elaborate_direct>Press the up arrow on the elevator. Use the lift
    and go to level 4. You find the button next
    to a yellow coloured marking.
  </elaborate_direct>
  <concise_indirect>There is an up arrow. The path continues on
    level 4.</concise_indirect>
  <elaborate_indirect>There is an up arrow on the elevator. The path
    continues on level 4. You find the button next
    to a yellow coloured marking.
  </elaborate_indirect>
</pathelement>
<pathelement id='13' nextElement='14'>
  <concise_direct>exit the lift and turn right following the corridor
    </concise_direct>
  <elaborate_direct>exit the lift, turn right and take the furthest
    right corridor. Follow the corridor around the
    snake bend.
  </elaborate_direct>
  <concise_indirect>after leaving the lift, your next turn is to the
    right. the path continues down the corridor
  </concise_indirect>
  <elaborate_indirect>after leaving the lift, your next turn is to
    the right and your way leads you to the
    furthest right corridor. The path continues down the corridor
    around the snake bend
  </elaborate_indirect>
</pathelement>
<pathelement id='14' nextElement='15'>
  <concise_direct>pass through the glass door</concise_direct>
  <elaborate_direct>at the end of the corridor, pass through the
```

```

        glass door with orange walls on the other side.
    </elaborate_direct>
    <concise_indirect>Find the glass door</concise_indirect>
    <elaborate_indirect>find the glass door with orange walls on the
        other side at the end of the corridor
    </elaborate_indirect>
</pathelement>
<pathelement id='15' nextElement='16'>
    <concise_direct>go up the stairs</concise_direct>
    <elaborate_direct>go up the stairs at the other side of the glass
        door</elaborate_direct>
    <concise_indirect>There are Stairs</concise_indirect>
    <elaborate_indirect>there are stairs at the other side of the
        glass door</elaborate_indirect>
</pathelement>
<pathelement id='16' nextElement=''>
    <concise_direct>Go through the door and continue straight until
        you reach room number 503</concise_direct>
    <elaborate_direct>At the top of the stairs, go through the door on
        the left hand side. Continue straight until
        you reach room number 503 which is the second door on the left
    </elaborate_direct>
    <concise_indirect>You see a door. The path continues straight
        until you find room number 503</concise_indirect>
    <elaborate_indirect>At the top of the stairs, you see a door on
        the left hand side. the path continues straight
        until you find room number 503 on the left hand side
    </elaborate_indirect>
</pathelement>
</path>

```

A.1.4. The Description Script of the third Route

```

<path>
    <pathelement id='1' nextElement='2'>
        <concise_direct>continue down the corridor</concise_direct>
        <elaborate_direct>continue down the corridor, passing the notice
            boards and fire extinguisher</elaborate_direct>
        <concise_indirect>the path continues down the corridor</
            concise_indirect>
        <elaborate_indirect>the path continues down the corridor, passing
            the notice boards and fire extinguisher
        </elaborate_indirect>
    </pathelement>

```

A. Additional Material

```
<pathelement id='2' nextElement='3'>
  <concise_direct>pass through the double doors and go up the stairs
  </concise_direct>
  <elaborate_direct>Pass through the set of double doors at the end
    of the corridor and go up the stairs.
  </elaborate_direct>
  <concise_indirect>Find the stairs behind the double doors.</
    concise_indirect>
  <elaborate_indirect>Find the stairs behind the double doors at the
    end of the corridor</elaborate_indirect>
</pathelement>
<pathelement id='3' nextElement='4'>
  <concise_direct>Turn right.</concise_direct>
  <elaborate_direct>At the top of the stairs turn right.</
    elaborate_direct>
  <concise_indirect>your next turn is to the right</concise_indirect
    >
  <elaborate_indirect>your next turn is to the right at the top of
    the stairs.</elaborate_indirect>
</pathelement>
<pathelement id='4' nextElement='5'>
  <concise_direct>go through the double doors</concise_direct>
  <elaborate_direct>go through the double doors with the big windows
    on the other side.</elaborate_direct>
  <concise_indirect>there are double doors</concise_indirect>
  <elaborate_indirect>there are double doors with big windows on the
    other side.</elaborate_indirect>
</pathelement>
<pathelement id='5' nextElement='6'>
  <concise_direct>continue down the corridor and follow it turning
    right.</concise_direct>
  <elaborate_direct>Continue down the corridor with the blue windows
    on the left-hand side. Pass the white lockers
    and turn right at the end of the corridor
  </elaborate_direct>
  <concise_indirect>the path continues down the corridor with the
    next section starting on your right
  </concise_indirect>
  <elaborate_indirect>The path continues down the corridor with the
    blue windows on the left-hand side. The next
    section starts on your right at the end of the corridor after
    passing the white lockers
  </elaborate_indirect>
```



```

</pathelement>
<pathelement id='6' nextElement='7'>
  <concise_direct>Turn left.</concise_direct>
  <elaborate_direct>walk forward to the green wall and then turn
    left.</elaborate_direct>
  <concise_indirect>your next turn is to the left</concise_indirect>
  <elaborate_indirect>your next turn is to the left when you see a
    green wall</elaborate_indirect>
</pathelement>
<pathelement id='7' nextElement='8'>
  <concise_direct>Turn right towards a green wall.</concise_direct>
  <elaborate_direct>Follow the corridor and turn right towards
    another green wall with the number 2 on it.
  </elaborate_direct>
  <concise_indirect>you see a green wall after turning right</
    concise_indirect>
  <elaborate_indirect>you see a green wall with the number 2 on it
    after following the corridor and turning right
  </elaborate_indirect>
</pathelement>
<pathelement id='8' nextElement='9'>
  <concise_direct>Turn right</concise_direct>
  <elaborate_direct>Turn right when you reach the wall</
    elaborate_direct>
  <concise_indirect>the path continues to the right</
    concise_indirect>
  <elaborate_indirect>when reaching the wall, the path continues to
    the right</elaborate_indirect>
</pathelement>
<pathelement id='9' nextElement='10'>
  <concise_direct>go towards the wall with the number 2 on it.</
    concise_direct>
  <elaborate_direct>Continue down the corridor and go towards the
    wall with the giant number 2 on it.
  </elaborate_direct>
  <concise_indirect>find the wall with the number 2 on it.</
    concise_indirect>
  <elaborate_indirect>find the wall with the giant number 2 on it at
    the other end of the corridor
  </elaborate_indirect>
</pathelement>
<pathelement id='10' nextElement='11'>
  <concise_direct>pass through two sets of double doors</

```

A. Additional Material

```
    concise_direct>
<elaborate_direct>Follow the corridor and pass through the two
    sets of big glass double doors</elaborate_direct>
<concise_indirect>There are two sets of double doors</
    concise_indirect>
<elaborate_indirect>there are two sets of big glass double doors
    after following the corridor
</elaborate_indirect>
</pathelement>
<pathelement id='11' nextElement='12'>
    <concise_direct>Turn right and go to the lift</concise_direct>
    <elaborate_direct>Turn right at the staircase and go to the lift
        with the green doors.</elaborate_direct>
    <concise_indirect>your next turn is to the right. Find the lift.</
        concise_indirect>
    <elaborate_indirect>your next turn is to the right when you see
        the staircase. Find the lift with the green
        doors.
    </elaborate_indirect>
</pathelement>
<pathelement id='12' nextElement='13'>
    <concise_direct>press the up arrow and go to level 4</
        concise_direct>
    <elaborate_direct>Press the up arrow on the elevator. Use the lift
        and go to level 4. You find the button next
        to a yellow coloured marking.
    </elaborate_direct>
    <concise_indirect>There is an up arrow. The path continues on
        level 4.</concise_indirect>
    <elaborate_indirect>There is an up arrow on the elvator. The path
        continues on level 4. You find the button next
        to a yellow coloured marking.
    </elaborate_indirect>
</pathelement>
<pathelement id='13' nextElement='14'>
    <concise_direct>exit the lift and turn right follwing the corridor
        </concise_direct>
    <elaborate_direct>exit the lift, turn right and take the furthest
        right corridor. Follow the corridor around the
        snake bend.
    </elaborate_direct>
    <concise_indirect>after leaving the lift, your next turn is to the
        right. the path continues down the corridor
```

```

    </concise_indirect>
    <elaborate_indirect>after leaving the lift, your next turn is to
        the right and your way leads your to the
        furthest right corridor. The path continues down the corridor
        around the snake bend
    </elaborate_indirect>
</pathelement>
<pathelement id='14' nextElement='15'>
    <concise_direct>pass though the glass door</concise_direct>
    <elaborate_direct>at the end of the corridor, pass though the
        glass door with orange walls on the other side.
    </elaborate_direct>
    <concise_indirect>Find the glass door</concise_indirect>
    <elaborate_indirect>find the glass door with orange walls on the
        other side at the end of the corridor
    </elaborate_indirect>
</pathelement>
<pathelement id='15' nextElement='16'>
    <concise_direct>go up the stairs</concise_direct>
    <elaborate_direct>go up the stairs at the other side of the glass
        door</elaborate_direct>
    <concise_indirect>There are Stairs</concise_indirect>
    <elaborate_indirect>there are stairs at the other side of the
        glass door</elaborate_indirect>
</pathelement>
<pathelement id='16' nextElement=''>
    <concise_direct>Go through the door and continue straight until
        you reach room number 503</concise_direct>
    <elaborate_direct>At the top of the stairs, go through the door on
        the left hand side. Continue straight until
        you reach room number 503 which is the second door on the left
    </elaborate_direct>
    <concise_indirect>You see a door. The path continues straight
        until you find room number 503</concise_indirect>
    <elaborate_indirect>At the top of the stairs, you see a door on
        the left hand side. the path continues straight
        until you find room number 503 on the left hand side
    </elaborate_indirect>
</pathelement>
</path>

```

A.2. Additional Evaluation Data

	indirect	direct	total
concise	8	3	11
elaborate	5	14	19
total	13	17	30

Table A.1.: The distribution of the communication styles of the user measured within the small talk dialogue.

A.2. Additional Evaluation Data

Cat.	Statement	M	σ	$M(p)$	$M(np)$	D	α
SRA	The system is accurate	3.67	1.15	3.87	3.57	0.30	0.2360
	The system is unreliable	2.16	1.14	1.97	2.25	0.28	0.2520
	The interaction with the system is unpredictable	1.92	1.01	1.80	1.98	0.18	0.3970
	The system didn't always do what I wanted	2.64	1.18	2.70	2.62	0.08	0.7900
	The system didn't always do what I expected	2.53	1.14	2.27	2.67	0.40	0.1060
	The system is dependable.	3.74	1.07	3.83	3.70	0.13	0.6960
	The system makes few errors.	3.57	1.26	3.63	3.53	0.10	0.7400
	The interaction with the system is consistent.	4.08	0.89	4.27	3.98	0.29	0.1640
L	The interaction with the system is efficient.	3.47	1.23	3.60	3.40	0.20	0.4420
	The system is useful.	4.03	1.10	4.20	3.95	0.25	0.1700
	The system is pleasant.	3.93	1.12	4.00	3.90	0.10	0.8710
	The system is friendly.	4.10	1.08	4.17	4.07	0.10	0.7370
	I enjoyed using the system.	4.06	1.25	4.13	4.02	0.11	0.7140
	It is clear how to speak to the system.	3.96	0.91	3.90	3.98	0.08	0.8380
	It is easy to learn to use the system.	4.47	0.80	4.50	4.45	0.05	0.8550
	I would use this system.	3.48	1.42	3.50	3.47	0.03	0.9930
CD	I felt in control of the interaction with the system.	4.01	0.97	4.13	3.95	0.18	0.4410
	I was able to fix errors easily.	3.63	1.05	3.83	3.53	0.30	0.1950
	I felt confident using the system.	4.00	1.13	4.13	3.93	0.20	0.4090
	I felt tense using the system.	1.94	1.06	1.90	1.97	0.07	0.5530
	I felt calm using the system.	4.02	1.19	4.27	3.90	0.37	0.1040
A	A high level of concentration is required when using the system	2.40	1.08	2.23	2.48	0.25	0.3110
	The system is easy to use.	4.21	0.98	4.27	4.18	0.09	0.9780
	The interaction with the system is repetitive.	3.21	1.29	3.13	3.25	0.12	0.7350
	The interaction with the system is boring.	2.12	1.16	2.23	2.07	0.16	0.4590
	The interaction with the system is irritating.	2.21	1.23	2.33	2.15	0.18	0.5250
H	The interaction with the system is frustrating.	1.93	1.16	1.77	2.02	0.25	0.4200
	The system is too inflexible.	2.64	1.15	2.67	2.63	0.04	0.7900
	I sometimes wondered if I was using the right word.	3.24	1.24	3.03	3.35	0.32	0.2610
	I always knew what to say to the system.	3.62	1.06	3.77	3.55	0.22	0.3280
SD	I was not always sure what the system was doing.	2.72	1.29	2.60	2.78	0.18	0.4670
	It is easy to lose track of where you are in an interaction with the system.	2.00	1.06	1.87	2.07	0.20	0.2640
	The interaction with the system is fast.	3.33	1.18	3.27	3.37	0.10	0.7380
	The system responded too slowly.	2.66	1.16	2.60	2.68	0.08	0.7240
SB	The system seemed polite.	4.29	1.01	4.33	4.27	0.06	0.8740
	The system seemed courteous.	3.96	1.13	4.03	3.92	0.11	0.7350
	The system seemed friendly.	4.18	1.08	4.27	4.13	0.14	0.7220
	The system seemed professional in its speaking style.	4.08	1.10	4.10	4.07	0.03	0.8470
SCS	The messages repeated themselves.	2.72	1.27	2.63	2.77	0.14	0.6530
	The system gave me more details than I needed.	1.88	0.97	1.90	1.87	0.03	0.8150
	The system spoke too much.	1.80	1.01	1.63	1.88	0.25	0.3160
	I felt that I had to wait too long for the system to stop talking in order to answer.	1.89	1.06	1.77	1.95	0.18	0.4410
	The system provided the right amount of information.	3.66	1.18	3.97	3.50	0.47	0.0640
	I got the information I wanted.	3.98	1.10	4.13	3.90	0.23	0.3420
	I would have liked more details.	2.50	1.29	2.47	2.52	0.05	0.9680
	I always knew what the system wanted to tell me.	3.93	1.19	3.97	3.92	0.05	0.8880
	The system didn't answer my questions accurately.	2.13	1.11	2.03	2.18	0.15	0.4470

Table A.2.: The results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(p)$ is the mean of the ratings of all dialogues were declared preferred and $M(np)$ is the mean of the ratings of all dialogues which were declared as not preferred. $D = |M(p) - M(np)|$ is the difference between $M(p)$ and $M(np)$, and α is the statistical significance level using the Mann-Whitney-U-Test.

A. Additional Material

Cat.	Statement	M	σ	$M(k_f, f)$	$M(k_f, r)$	$M(k_f, a)$	α
SRA	The system is accurate	3.67	1.15	3.73	3.67	3.73	0.9720
	The system is unreliable	2.16	1.14	2.00	2.60	2.13	0.2290
	The interaction with the system is unpredictable	1.92	1.01	2.20	2.07	1.67	0.4150
	The system didn't always do what I wanted	2.64	1.18	2.67	2.80	2.60	0.8980
	The system didn't always do what I expected	2.53	1.14	2.60	2.67	2.47	0.8810
	The system is dependable.	3.74	1.07	3.73	3.67	3.87	0.8790
	The system makes few errors.	3.57	1.26	3.07	3.33	3.80	0.2140
	The interaction with the system is consistent.	4.08	0.89	4.13	3.73	4.07	0.2990
	The interaction with the system is efficient.	3.47	1.23	3.27	3.40	3.67	0.5390
L	The system is useful.	4.03	1.10	3.80	3.80	4.07	0.7360
	The system is pleasant.	3.93	1.12	3.67	3.73	4.00	0.6650
	The system is friendly.	4.10	1.08	4.07	4.13	4.27	0.8030
	I enjoyed using the system.	4.06	1.25	3.73	3.73	4.00	0.7970
	It is clear how to speak to the system.	3.96	0.91	4.07	4.13	4.20	0.8450
	It is easy to learn to use the system.	4.47	0.80	4.60	4.40	4.53	0.9320
	I would use this system.	3.48	1.42	3.20	3.27	3.20	0.9600
	I felt in control of the interaction with the system.	4.01	0.97	3.80	3.93	3.87	0.9020
	I was able to fix errors easily.	3.63	1.05	3.53	3.47	3.47	0.9620
CD	I felt confident using the system.	4.00	1.13	4.13	3.93	4.00	0.8960
	I felt tense using the system.	1.94	1.06	2.00	2.07	2.00	0.9930
	I felt calm using the system.	4.02	1.19	4.07	4.00	4.07	0.9620
	A high level of concentration is required when using the system	2.40	1.08	2.53	2.47	2.47	0.9720
	The system is easy to use.	4.21	0.98	4.20	4.27	4.20	0.9600
A	The interaction with the system is repetitive.	3.21	1.29	3.47	3.40	3.33	0.9770
	The interaction with the system is boring.	2.12	1.16	2.20	2.40	2.07	0.8200
	The interaction with the system is irritating.	2.21	1.23	2.60	2.40	2.20	0.7220
	The interaction with the system is frustrating.	1.93	1.16	2.33	2.07	1.87	0.6080
	The system is too inflexible.	2.64	1.15	2.80	3.07	2.93	0.8840
H	I sometimes wondered if I was using the right word.	3.24	1.24	2.80	3.20	3.07	0.5840
	I always knew what to say to the system.	3.62	1.06	3.80	3.60	3.80	0.7740
	I was not always sure what the system was doing.	2.72	1.29	2.73	2.73	2.67	0.9910
	It is easy to lose track of where you are in an interaction with the system.	2.00	1.06	2.33	2.40	1.73	0.2270
SD	The interaction with the system is fast.	3.33	1.18	3.07	3.47	3.20	0.6550
	The system responded too slowly.	2.66	1.16	3.00	2.73	2.27	0.1830
SB	The system seemed polite.	4.29	1.01	4.13	4.27	4.33	0.8250
	The system seemed courteous.	3.96	1.13	3.87	3.93	3.87	0.9690
	The system seemed friendly.	4.18	1.08	4.07	4.13	4.27	0.7620
	The system seemed professional in its speaking style.	4.08	1.10	3.93	3.93	3.93	0.9920
SCS	The messages repeated themselves.	2.72	1.27	2.60	2.67	2.40	0.8000
	The system gave me more details than I needed.	1.88	0.97	2.07	2.33	1.93	0.4650
	The system spoke too much.	1.80	1.01	1.67	2.33	1.80	0.2210
	I felt that I had to wait too long for the system to stop talking in order to answer.	1.89	1.06	1.80	2.27	1.60	0.2510
	The system provided the right amount of information.	3.66	1.18	3.60	3.53	3.73	0.9470
	I got the information I wanted.	3.98	1.10	3.93	3.73	3.87	0.8580
	I would have liked more details.	2.50	1.29	2.53	2.27	2.53	0.7770
	I always knew what the system wanted to tell me.	3.93	1.19	3.60	3.93	3.60	0.7540
	The system didn't answer my questions accurately.	2.13	1.11	2.67	2.07	1.80	0.0590*

Table A.3.: The results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(k_f, x)$ is the mean of the ratings of all dialogues with full knowledge, where $f = \text{fixed}$, $r = \text{random}$ and $a = \text{adaptive}$. α is the statistical significance level using the Kruskal-Wallis-Test. The differences that are statistically significant ($\alpha < 0.05$) are marked (*).

A.2. Additional Evaluation Data

Cat.	Statement	M	σ	$M(k_p, f)$	$M(k_p, r)$	$M(k_p, a)$	α
SRA	The system is accurate	3.67	1.15	3.53	3.80	3.53	0.7380
	The system is unreliable	2.16	1.14	1.93	2.00	2.27	0.6770
	The interaction with the system is unpredictable	1.92	1.01	1.93	1.60	2.07	0.4640
	The system didn't always do what I wanted	2.64	1.18	2.47	2.47	2.87	0.6310
	The system didn't always do what I expected	2.53	1.14	2.73	2.40	2.33	0.8110
	The system is dependable.	3.74	1.07	3.80	3.67	3.73	0.8800
	The system makes few errors.	3.57	1.26	3.53	4.00	3.67	0.6380
	The interaction with the system is consistent.	4.08	0.89	4.13	4.33	4.07	0.6530
	The interaction with the system is efficient.	3.47	1.23	3.53	3.40	3.53	0.8470
L	The system is useful.	4.03	1.10	4.07	4.33	4.13	0.7640
	The system is pleasant.	3.93	1.12	4.07	4.20	3.93	0.9980
	The system is friendly.	4.10	1.08	3.93	4.27	3.93	0.7150
	I enjoyed using the system.	4.06	1.25	4.27	4.53	4.07	0.6510
	It is clear how to speak to the system.	3.96	0.91	3.80	3.73	3.80	0.9900
	It is easy to learn to use the system.	4.47	0.80	4.47	4.47	4.33	0.9350
	I would use this system.	3.48	1.42	3.67	3.80	3.73	0.9990
	I felt in control of the interaction with the system.	4.01	0.97	4.27	4.13	4.07	0.8140
	I was able to fix errors easily.	3.63	1.05	3.60	3.93	3.80	0.6840
CD	I felt confident using the system.	4.00	1.13	4.07	3.93	3.93	0.6200
	I felt tense using the system.	1.94	1.06	1.87	1.67	2.07	0.8820
	I felt calm using the system.	4.02	1.19	4.07	4.13	3.80	0.8850
	A high level of concentration is required when using the system	2.40	1.08	2.40	2.13	2.40	0.7800
	The system is easy to use.	4.21	0.98	4.27	4.27	4.07	0.8950
A	The interaction with the system is repetitive.	3.21	1.29	3.27	3.13	2.67	0.3820
	The interaction with the system is boring.	2.12	1.16	2.13	1.87	2.07	0.8420
	The interaction with the system is irritating.	2.21	1.23	2.13	1.80	2.13	0.6590
	The interaction with the system is frustrating.	1.93	1.16	1.93	1.47	1.93	0.6990
	The system is too inflexible.	2.64	1.15	2.00	2.47	2.60	0.2560
H	I sometimes wondered if I was using the right word.	3.24	1.24	3.47	3.53	3.40	0.9480
	I always knew what to say to the system.	3.62	1.06	3.53	3.67	3.33	0.7190
	I was not always sure what the system was doing.	2.72	1.29	2.80	2.67	2.73	0.9730
	It is easy to lose track of where you are in an interaction with the system.	2.00	1.06	1.80	1.87	1.87	0.9940
SD	The interaction with the system is fast.	3.33	1.18	3.40	3.53	3.33	0.8680
	The system responded too slowly.	2.66	1.16	2.67	2.53	2.73	0.8900
SB	The system seemed polite.	4.29	1.01	4.20	4.53	4.27	0.6850
	The system seemed courteous.	3.96	1.13	4.07	4.00	4.00	0.9380
	The system seemed friendly.	4.18	1.08	4.13	4.33	4.13	0.9680
	The system seemed professional in its speaking style.	4.08	1.10	4.13	4.13	4.40	0.7510
SCS	The messages repeated themselves.	2.72	1.27	3.13	2.87	2.67	0.6350
	The system gave me more details than I needed.	1.88	0.97	1.40	1.53	2.00	0.1170
	The system spoke too much.	1.80	1.01	1.60	1.53	1.87	0.9740
	I felt that I had to wait too long for the system to stop talking in order to answer.	1.89	1.06	1.47	2.00	2.20	0.1230
	The system provided the right amount of information.	3.66	1.18	3.87	3.87	3.33	0.3430
	I got the information I wanted.	3.98	1.10	4.13	3.93	4.27	0.5330
	I would have liked more details.	2.50	1.29	2.07	2.60	3.00	0.1470
	I always knew what the system wanted to tell me.	3.93	1.19	4.13	4.40	3.93	0.8110
	The system didn't answer my questions accurately.	2.13	1.11	2.07	1.80	2.40	0.5650

Table A.4.: The results of the participants' ratings, whereby M is the mean and σ is the standard deviation of all ratings, $M(k_f, x)$ is the mean of the ratings of all dialogues with partial knowledge, where $f = \text{fixed}$, $r = \text{random}$ and $a = \text{adaptive}$. α is the statistical significance level using the Kruskal-Wallis-Test.

A. Additional Material

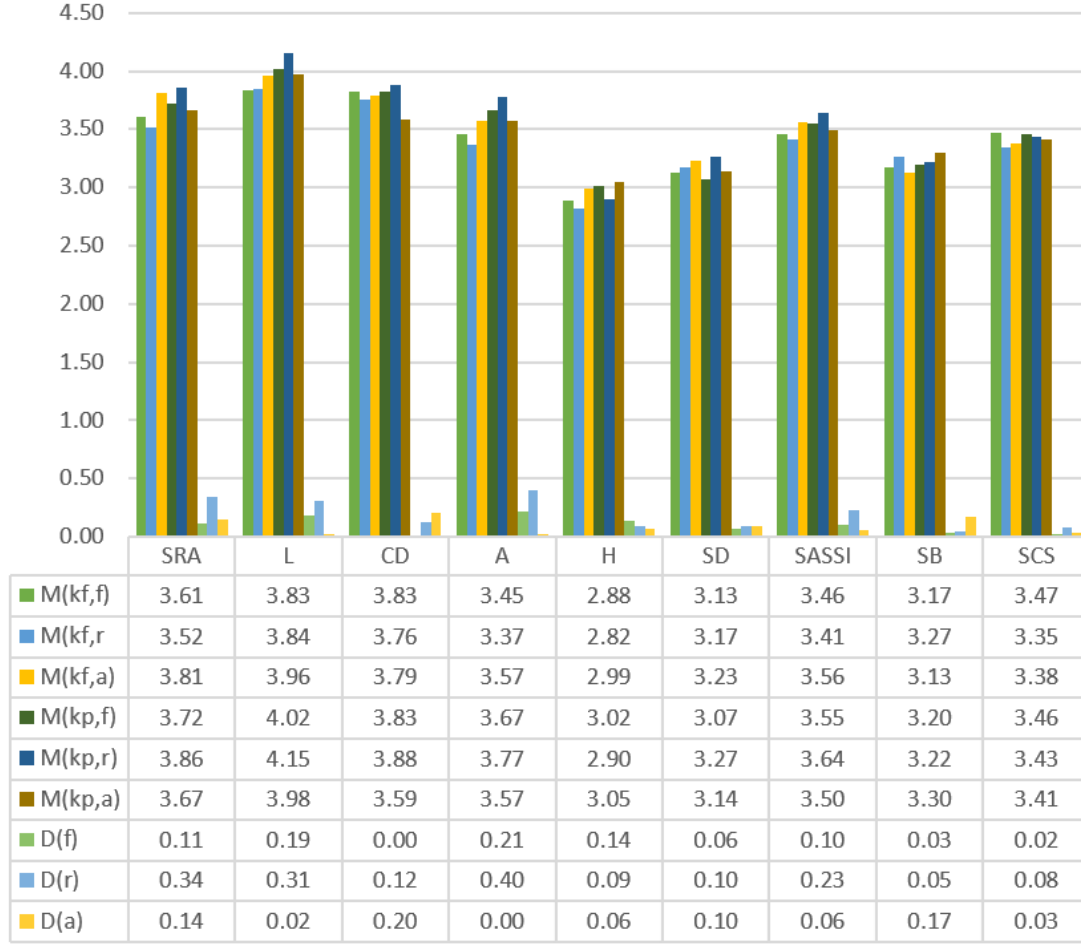


Table A.5.: The results of the participants' ratings grouped by the different categories of the questionnaire, whereby the rating scale has been inverted for the negatively formulated statements so that the optimal rating is five for every statement. $M(k_f, x)$ is the mean of the ratings of all dialogues with full knowledge and the communication style strategy x ($f = fixed, r = random, a = adaptive$) and $M(k_p, x)$ is the mean of the ratings of all dialogues with partial knowledge and the communication style strategy x ($f = fixed, r = random, a = adaptive$), $D(f) = M(k_f, f) - M(k_p, f)$, $D(r) = M(k_f, r) - M(k_p, r)$ and $D(a) = M(k_f, a) - M(k_p, a)$ are the differences between the mean values.