Hey You! Let's Talk. Dialogue-Initiatives Revisited for Wayfinding Instructions

Pauline Krieger[1][0000−0002−7808−0448], Markus Kattenbeck[2][0000−0001−6028−0428],
Bernd Ludwig[1][0000−0003−3599−2081], Johannes Helmbrecht[1], and
Ioannis Giannopoulos[2][0000−0002−2556−5230]

[1] University of Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany
`pauline.krieger@student.ur.de`,
`bernd.ludwig@ur.de`,`johannes.helmbrecht@ur.de`

[2] TU Wien, Gußhausstraße 27-29, 1040 Wien, Austria
`markus.kattenbeck@tuwien.ac.at`,
`ioannis.giannopoulos@tuwien.ac.at`

Abstract. This paper presents Human-Computer Interaction design guidelines for interactive wayfinding assistance systems which provide on-line route instructions. These design suggestions are based on a corpus of human-to-human, on-line, landmark-based route instructions in German language which were gathered by means of an in-situ study involving pairs of participants. Based on the description of this collection, which is made publicly available, an in-depth analysis of the corpus is presented: This analysis reveals the importance of establishing Common Ground through existential-presentative constructions which have, up until now, not been taken into account in presenting route instructions to users of pedestrian navigation systems. These syntactical constructs provide the empirical ground for two important design suggestions: Systems should, first, ask for explicit feedback whether a salient object is recognised by users before referring to this object in a route instruction. Second, a mode of negotiating Common Ground once it was lost should be implemented, which can be initiated by the user. The results reveal the importance of the state-tracking capabilities of wayfinding assistance systems.

Keywords: Existential-presentative Constructions, Human Wayfinding, Corpus, Spoken Language, Route Instructions, Common Ground

1 Introduction

Despite the complexity of the task, wayfinding, i.e. deliberately getting from a starting point to a destination in a familiar or unfamiliar environment [see e.g. 1, 2], is successfully achieved by humans on an everyday basis. Research in systems providing assistance in solving this task and decreasing its complexity has raised considerable interest [see e.g. 3]. Among these efforts, route instructions per se, their calculation and their presentation have seen much interest in research on wayfinding assistance systems (see Section 2). The empirical foundation of the phrasing used for these instructions has, however, only raised interest in recent years [see 4]. With respect to the design of wayfinding assistance

systems, human-to-human, real-world route instructions are particularly worthwhile to study from a linguistics point of view [see 5] for two reasons: First, the linguistic interactions can be mimicked by these systems in order to increase the likeliness of successful wayfinding. Second, the linguistic properties give an insight into how systems should act to resolve misunderstandings and, thereby, (re-)establish so-called Common Ground (CG). Up until now, however, we lack sufficient knowledge about how elements of discourse relate to wayfinding success [see 6]. In this paper, we try to add to this knowledge by

1. presenting the POPRIS-corpus of turn-by-turn route instructions given over the phone; this corpus was collected in an in-situ experiment in an environment participants were unfamiliar with;
2. providing evidence for the importance of CG and different modes of establishing;
3. deriving suggestions for Human-Computer Interaction (HCI) design for (re-)establishing CG in wayfinding assistance systems.

2   Related Work

In-line with the focus of our paper we review three strands of related work. An overview of existing corpora on route instructions will be used to provide a rationale why collecting a new corpus is reasonable. We continue with existing literature from a linguistics perspective. Finally, we shed light on work on route instruction generation in the domain of wayfinding assistance systems.

2.1   Corpora of Route Instructions

Although studies on the elements of route instructions regularly collect a corpus (i.e. "[a] collection of linguistic data, either compiled as written texts or as a transcription of recorded speech" [7]), only few corpora are made publicly available. Recently, Liu, Tree, and Walker [8] present a corpus of over the phone in-situ conversations during which the sender S is required to guide the receiver R to several different artworks based on maps. Similar to our German-language POPRIS-corpus, this corpus encompasses transcriptions of full English language conversations. Götze and Boye [9] present and describe the SPACEREF corpus which consists of descriptions of actions and the spatial environment of a pedestrian while walking a predefined route in a Swedish town. Tenbrink et al. [6] present a corpus full of linguistic expressions of spatial relations in German; these, however, were collected in a "referential communication task that involved furnishing a dolls' house" [6, p. 2]. Other dialogue corpora with spatial expressions have been compiled during experimental tasks to explain routes on maps in order to study cognitive reasoning during complex tasks. The HCRC Map Task Corpus [10] is a well-known example for this type of collection. Corpora have also been used to compare L1- and L2-learners and their capabilities to generate spatial expressions [see e.g. 11, 12, 13].

Taken together, several of these corpora include route instructions. Compared to these, the contribution of our POPRIS-corpus is twofold: 1) It contains the full dialogues

between pairs of people unfamiliar with the environment during the whole experiment; 2) data was gathered by means of an in-situ experiment during which on-line, turn-by-turn route instructions over the phone were collected from participants, who were not allowed to use any aids to perform the task. Given this setting, these dialogues are expected to yield significant insights into referential, real-world collection of utterances suitable to increase HCI dialogue structures.

2.2 Elements of route instructions

The elements of route instructions have seen a long-lasting interest among a range of disciplines, including linguistics [see 14, for an extensive overview], psychology, geography [see 15, pp. 65–67 for early references] as well as AI and cognitive science. Consequently, the amount of literature into route instructions is vast. This overview can, therefore, only mention highlights on these different strands; we, therefore, focus on the Lingustics and GIScience perspective.

Linguistics Detailed analyses of elements of route instructions have been conducted throughout the years, essentially starting in the late 1970s/early 1980s. Wunderlich [16] authored a highly influential publication for the assessment of German language route instructions. He provides a discourse-theoretic analysis of these instructions and, thereby, identifies different discourse markers for different segments of a route. Klein [17] uses a corpus of route instructions given to a receiver prior to navigation in order to analyse deictic elements. Despite the fact that Klein neither focuses nor points to these, his examples resemble topic-constructions based on existential-presentative constructions (EPCs) [18] as found in our corpus.

   More recent analyses have, for instance, focused on differences between males and females in in-situ collected directions given to car drivers [19]; others have investigated lexical choice and referring expressions in over-the-phone instructions [20]. Generally speaking, linguistic accounts of route instructions have recently regained momentum through Cognitive Discourse Analysis [CODA, 4]. This method is particularly suitable to examine how decision points are linguistically represented (e.g. in terms of syntactic position). It is, therefore, of particular interest to our study, as we provide evidence that EPCs are syntactic patterns which are used to mark decision points by means of landmarks (see Section 5).

GIScience The importance of landmarks for spatial cognition and, consequently, their prevalence in route instructions has been studied frequently. Lovelace, Hegarty, and Montello [15] introduce different types and roles of landmarks and present evidence that the quality of route instructions is a function of the types of landmarks and its combinations. Since then, numerous attempts have been made to model [see e.g. 21, 22, 3, 23, 24] and measure the salience of objects [see e.g. 25, 26, 27, 28] as this is a prerequisite to enable references to landmarks in route instructions.

   In addition to that, considerable effort has been put into creating and assessing systems which are suitable to provide route instructions to users based on landmarks. The adaptation of route instructions to the spatial context or user knowledge and preferences

in Artificial Intelligence and Cognitive Science research has seen considerable interest as early as the mid-1990s: For example, the MOSES project [29, 30, 31] investigated route generation by taking the perception and changes of spatial layout during navigation into account. These efforts already resemble the problem of appropriate levels of granularity in route instructions which has been of particular importance in numerous studies. Tomko and Winter [32] present a recursive algorithm to determine route instructions which adapt their level of granularity to the distance to the destination. Based on wayfinding choreme theory, Richter and Klippel [33] develop a hierarchy of route direction elements. Richter, Tomko, and Winter [34] suggest a dialogue-based solution in which users can request more details to solve the issue of adapting the instruction's granularity to a user's knowledge. Finally, cognitively ergonomic route instructions based on chunking [35, 36] have been a milestone in research on adapting route instructions to user knowledge. Researchers have, moreover, investigated suitable means to include landmarks in route instructions from a computational linguistics perspective [see e.g. 37]. Disambiguation of landmarks has been identified as a major issue in these efforts; the use of EPCs as topic construction implies this disambiguation (see Section 5.1). It is important to note that the corpus we present may be equally useful for each of these disciplines. Having said this, our analysis focuses on the discourse, sentence and dialogue structures in human-to-human conversations in order to derive HCI-guidelines for wayfinding assistance systems.

3 Experimental Design

In this section we provide details for the experimental design based on which we collected the corpus of route instructions. The whole experimental design is based on the goal to study the discourse which leads to successful route instructions in a real-world wayfinding situation.

3.1 Routes

The experiments took place in different parts of a small European town (Regensburg) of Roman decent, its suburbs or close-by villages depending on which area potential participants were not familiar with (see below for further explanations). In order to find comparable routes across experimental areas (see [3] for a model on wayfinding decision situations), we made use of the intersections framework introduced in [38]. The chosen routes (see Section 4.2) reflect the relative frequency of the 3- and 4-way intersections (see Table 1), which were the two topmost intersection types.

Table 1. Characteristics of the routes; *Br* denotes branches, *Len.* denotes the length of route, *Dur.* denotes the walking duration according to Google Maps, PDP means potential decision point, TDP means true decision point, *%(3)* and *%(4)* describe the relative frequencies of 3- and 4-way intersections included in the route, *Norm (3)* and *Norm (4)* display the relative frequencies of 3- and 4-way intersections in the whole area of Regensburg as reference values. Due to the spatial configuration of area 4, however, it was impossible to include 4-way intersections in these routes.

Route	3 Br.	%(3)	Norm (3)	4 Br.	%(4)	Norm (4)	Len. [km]	Dur. [min]	PDPs	TDPs
1.1	21	0.88	0.81	2	0.08	0.19	1.5	18	18	6
1.2	23	0.96	0.81	1	0.04	0.19	1.2	16	16	8
2	16	0.84	0.81	3	0.16	0.19	1.6	20	10	9
3	14	0.82	0.81	3	0.18	0.19	1.1	13	9	8
4.1	14	1	0.81	0	0	0.19	1.2	15	8	6
4.2	11	1	0.81	0	0	0.19	1.1	12	1	10
5.1	10	0.67	0.81	5	0.33	0.19	1.2	14	5	10
5.2	10	0.67	0.81	4	0.27	0.19	1.1	14	9	6
6	21	0.84	0.81	4	0.16	0.19	1.2	16	14	11
7	18	0.95	0.81	1	0.05	0.19	1.6	19	8	11
8.1	24	0.77	0.81	7	0.23	0.19	1.2	18	25	6
8.2	21	0.81	0.81	5	0.19	0.19	1.7	22	17	9
8.3	23	0.79	0.81	6	0.21	0.19	1.7	21	16	13
9	13	0.81	0.81	3	0.19	0.19	1.5	20	9	7
10.1	28	0.85	0.81	5	0.15	0.19	1.4	17	23	10
10.2	18	0.82	0.81	4	0.18	0.19	1.3	16	17	5
11	23	0.74	0.81	7	0.23	0.19	1.3	16	19	12
12	28	0.97	0.81	1	0.03	0.19	1.2	15	19	10
13	8	0.67	0.81	4	0.33	0.19	1.4	18	5	7
14	13	0.65	0.81	7	0.35	0.19	1.5	19	10	10

3.2 Finding pairs of participants

Potential participants were acquired through leaflets distributed at the local university. Any person who expressed interest, was required to fill in an online survey on sense of direction (FRS, [39]), the Big Five Personality Traits (BFI-60, [40]), demographic data (age, gender) and their spatial knowledge about several regions: For each of 14 different outdoor environments, which were clearly marked as polygons on a map, participants were required to indicate whether they have ever been there before and, if yes, how good they knew this area. The goal of this procedure was to find environments with which participants were not familiar. This means, all experiments were conducted in an environment participants were not familiar with, in order to avoid potential bias from familiarity (see [41] for evidence on the impact familiarity has on the way route instructions are given). Once a pool of potential participants was found, pairs of participants were randomly chosen, roles (sender/receiver) were assigned randomly and – if a pair

of participants indicated to be unfamiliar with more than one experimental area – the area was chosen randomly, too.

3.3 Procedure

In order to ensure no interactions between sender and receiver, both were required to arrive separately at the starting point of their assigned route. The experimental procedure comprised five steps:

1. A student experimenter and the sender met at a bus stop close to the starting point of the assigned route.
2. The sender was instructed that the experimenter would guide her/him on a route and that s/he will later be required to give route instructions to the receiver over the phone and that these should enable the receiver to follow the route exactly.
3. The experimenter guided the sender along the predefined route, avoiding conversations as much as possible in order to prevent potential bias.
4. At the destination, senders were asked to wait until the receiver would call and reminded about their task.
5. The experimenter met the receiver at the starting point and instructed her/him that the sender would guide her/him over the phone on a specific route and that s/he is required to follow this route exactly. The experimenter did not intervene, but followed the receiver silently and noted down any wrong turns (if any). The conversations between sender and receiver were audio recorded.

4 POPRIS-Corpus

4.1 An overview of corpus data

The POPRIS-corpus consists of the full transcripts of sixteen experiments ($N = 32$ participants) covering 3:22:32 hours of German language audio material in total. Audio recordings were transcribed using the software f4transkript [see 42] in accordance with the HIAT [see 43] guidelines. In particular, the following spoken-language linguistic phenomena are part of the transcripts: pauses, overlap, hesitation, emphasis, repair, lengthening, sentence type, anacoluthon, linking and non-verbal information. The complete discourse between sender and receiver is included in the transcripts. As the full experiment was recorded and transcribed the transcripts cover also conversation between sender and receiver which is not related to the route guidance task itself; this yields a full picture of conversations occurring during this type of task potentially valuable for other disciplines. Out of 16 experiments one experiment was abandoned on request by the sender, who admittedly was unable to guide the receiver along the route. The remaining 15 experiments had a mean duration of 00:16:55 minutes ($MD = 00:16:44$, $MIN = 00:08:47$, $MAX = 00:45:45$). Overall, the corpus encompasses 203 instructions, out of which 176 were successful (i.e. a correction happened before the upcoming TDP was reached), and 27 failed. The instructions consist of 2,466 turns [counted according to 6, p. 15] in total, 1,172 of them given by senders and 1,294 issued by receivers. Seven out of the $N = 32$ participants were females, one person indicated *diverse* as gender, one pair

was female only and nine pairs were male only. In general, participants were recruited from a rather young age group ($M = 20.8$, $MIN = 17$, $MAX = 27$, $MED = 20$).

In addition to the transcripts, the corpus includes age and gender as demographic variables, the 19-item questionnaire *Fragebogen räumliche Strategien* (i.e. a German language sense of direction scale, [39]) and the 60 items of the BFI-2 [40]. Based on the norm data of the German population [44] the majority of the participants has generally high spatial abilities. These figures reflect the fact that the willingness to participate in a wayfinding experiment is increased for those who perceive their sense of direction as good. As far as the Big Five Personality Traits are concerned, participants show mainly (above) average values (see [40] for norm data) for agreeableness (f: 3.80, m: 3.64) and openness (f: 3.39, m: 3.28), (below) average values for conscientiousness (f: 3.72, m: 3.57) and heterogeneous values for neuroticism (f: 2.82, m: 2.66) and extraversion (f: 3.27, m: 3.13).

4.2 Data and Software Availability

The POPRIS-Corpus is publicly available through the `zenodo.org` platform (DOI: 10.5281/zenodo.3695744). It includes the transcripts, a QGIS-file containing all routes as well as age, gender, FRS and BFI-2 data for each of the participants.

5 Results

5.1 Important Definitions

The analysis is based on three important linguistic concepts, namely *Common Ground (CG) Content*, *Topic Construction* and *Existential-presentative constructions (EPCs)*. CG content can be seen as the shared knowledge of interlocutors, which consists of propositions and already introduced entities [see e.g. 45]. It is subject to continuous change throughout the conversation [see 46, p. 245]. Generally speaking, topic constructions consist of two parts, the topic and the comment: "The topic constituent identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the CG content" [45, p. 5]. *Existential-presentative constructions (EPCs)* [18] are one of several existing types of topic constructions. As landmarks are entities and references to these are known to be used in human-to-human route instructions, EPCs are an obvious mean of CG management [46] and will be used as a basis for our analysis.

5.2 Discourse Analysis Results

The importance of EPCs The whole corpus contains three different cases of how EPCs are used in instructions, two of which show important ways of repairing insufficient CG. The corpus data gathered provide evidence that EPCs are the most common syntactic way of introducing a new (i.e. up until now not mentioned) object, which stands out in the local environment. Most frequently, the EPC established a CG, which, in turn, resulted in a comprehensible route instruction. Sample Dialogue *Case 1* provides an example for this type of discourse:

Sample Dialogue Case 1: A dialogue example during which CG has been successfully established using an EPC.

Role	Dialogue
S:	Und ähm siehst du/ da is so n anderes Bushaltestellenschild. Das heißt „Am Bach". Das is, wenn du über die //Stra//ße gehst Richtung...
	And uh can you see/ there is another bus stop sign. It is called "Am Bach". When you cross the street it is in the direction...
R:	//Ja.//
	Yes.
S:	Genau. Und da //gehst// jetz einfach mal drauf zu.
	Right. And now you just go towards it.

In this case, a route instruction is split into three subsequent actions: The sender uses an EPC ("And uh can you see/ there is another bus stop sign") to introduce a new entity. The sender, thereby, expects the receiver to explicitly confirm that s/he can identify the object s/he referred to in the local environment. This is a mandatory step in order to ensure that both, sender and receiver, share a CG. When CG is achieved (confirmation by the receiver: "Yes"), the sender continues with the actual route instruction ("And now you just go towards it") – which is, in fact, the comment of the topic construction.

Occasionally (27 times) however, the POPRIS-corpus contains discourse segments which show a misunderstanding between sender and receiver. Two different variants exist, in both of which CG is not successfully established, but for different reasons. In variant 1, CG is not achieved, although both, sender and receiver, essentially share it. This means, the sender imagines the receiver to be at a particular location where the receiver in-fact is. Yet, their communication about the location is unsuccessful. Sample Dialogue *Case 2* provides an example:

Sample Dialogue Case 2: A dialogue example during which a misinterpretation (church vs. right-hand bend) occurs between sender and receiver and is repaired by introducing new salient objects and references to those which have already been part of the CG content.

Role	Dialogue
S:	Ja. (...) Dann einfach immer noch der Straße folgen (...) und dann müsste die unten so ne leichte Rechtskurve machen.
	Now continue along this street until you reach a slight right-hand bend.
R:	Ok. Da sind wir.
	Ok. We have arrived.
S:	Siehst die Rechtskurve? Genau und dann eben äh der Straße weiter in dieser Rechtskurve folgen.
	Do you see the right-hand bend? Now follow the street further along this bend.
R:	Also die Rechtskurve, in der das gelbe Haus drinsteht (...) oder erst da//nach//?
	So the right-hand bend where the yellow house is located (...) or thereafter?
S:	//Mh..// danach.
	Mh.. thereafter.
R:	O//k.//
	Okay.

S: //A//lso eigentlich einfach nur die/ die Straße weiter gradeaus gehen? ((lacht))·
So actually just follow the/ the street a little further? ((laughs))·

R: Ok, wir sind jetz an der Kirche. Oder so gut wie.
Okay, we have almost arrived at the church.

S: An was?
Where?

R: An der Kirche.
At the church.

S: An der Kirche ((lacht))· Ähm ich erinner' mich leider nich an ne Kirche ((lacht))· Mh //ja...//
At the church ((laughs))· Uhm unfortunately I don't remember a church ((laughs))· Mh yeah...

R: //Äh// ok. Also wir sind jetz grad an dem gelben Haus vorbei. Also //äh// (...) komisches gelb äh, ich hab' da vorher äh das andere Haus gemeint, also so gelb-grünlich äh...
Uh okay. So we have just passed the weird yellow-greenish house. I've talked about uh the other house before.

S: //Mhm.// (...) Mhm.
Mhm. (...) Mhm.

R: Genau dahinter (unv.) oder das Haus stand direkt in der Rechtskurve.
Right behind (imcompr.) or the house is located directly in the right-hand bend.

S: Genau, ja.
Right, yes.

R: Äh und jetz stehen wir da vorne vor ner Kreuzung mit einem äh ja äh/ vor einer Kreuzung, die jetz (unv.) is und hinter der Kreuzung (links) steht eine Kirche.
Uh and now we are standing in front of an intersection and behind the intersection (to the left) is a church.

S: Ah ja, mhm ok. (...) Äähm ja da geht's jetz so leicht rechts den Berg so n bisschen runter, oder? (...) Also man könnte so links gehen. Da geht's so hoch, glaub ich, und dann aber rechts.
Ah yes, mhm okay. (...) Uh yeah there is a hill which goes slightly down to the right, isn't there? (...) To the left you could go up the hill but you need to turn right.

R: Äh ja, also links/ äh also, wenn wir ein bisschen zurück gehen, dann geht's links den Berg rauf. Da steht auch ein äh Kreuz. Also dir//ekt hinter// dem gelben Haus, geht's links rauf.
Uh yes, so when we go back a little, I can go up the hill. There is also a cross. So immediately behind the yellow house I could go left and up the hill.

S: //Genau da// (...) Genau da aber nich, sondern einfach (in dieser) Rechtskurve da. //Wieder so n bisschen runter.//
Right there (...) Right but don't go there but instead just follow the right-hand bend. A bit down the street again.

R: //Ja ok, dann// wären wir an der Kreuzung vor der Kirche.
Yes okay then we would be at the intersection in front of the church.

S: Mhm und was habt ihr jetz für Optionen, wo du hingehst?
Mhm and what options do you have now?

R: Äh gradeaus, dann komm ich an der Kirche vorbei //äh// oder links den Berg rauf oder rechts den Berg runter.
Uh I could go straight and I will pass the church uh or I could turn left and go up the hill or turn right and go down the hill.

S: //Mhm.// (...) Rechts den Berg runter.
Mhm. (...) Right down the hill.

Role	Dialogue
R:	Ok. Also wieder in die Richtung, aus der wir gekommen s/ äh in der die Bushaltestelle liegt. *Okay. So again in the direction from which we came/ uh in which the bus stop lies?*
S:	Genau, ja. *Right, yes.*

The sender uses an EPC to introduce a feature of the street network ("right-hand bend") as a salient object. The receiver, however, does not confirm this immediately but introduces a salient object ("yellow house"), too. One reason might be, that the road network feature is less salient to R than the colour of the facade. From this turn on, sender and receiver swap their roles, i.e. R introduces salient objects using EPCs (e.g. "church") and expects the sender to confirm spatial knowledge about these. S, however, does not remember a church and CG is not achieved. The roles, consequently, remain swapped until CG is established.

In contrast to variant 1, CG cannot be established in variant 2 because it was not established as part of the preceding route instruction. Sample Dialogue *Case 3* provides an example for reestablishing CG. Please note: In the previous step, R had mistakenly confirmed the landmark ("silver bridges") which led S to assume that CG was established which it in fact was not.

Sample Dialogue Case 3: A dialogue example during which a misunderstanding ("silver bridges" were mistakenly confirmed) occurs between sender and receiver. Parts of the discourse which were not relevant for the repair are left out (marked by dots).

Role	Dialogue
S:	Gen//au// und irgendwann ähm auf der rechten Seite bevor die Häuser wieder so wirklich losgehen, äh kommt so n kleiner Fußgängerweg, (...) der dann //rechts// reingeht. //Genau.// *Right and at some point a small walkway will appear on your right side before the houses start again (...) The walkway makes a right turn. Right.*
...	...
S:	//A//ber ich glaube/ Genau und/ also da fließt auch auf jeden Fall der Bach dann sozusagen um so n (...) so ne Ecke. Also das sind glaub' ich so zwei Bächlein, die dann da kommen. *The rill makes a bend. So two rills will appear there.*
R:	Also gleich nach der Bushaltestelle hätt' ich dann links reingehen sollen oder wie? *So right after the bus stop I should have gone left, right?*
...	...
R:	Ja, also ich glaub' ich hab' jetzt den Weg gefunden, ((lacht))· den du meinst. //(unv.)// an der linken Seite sind dann so Gärten und so? *Yes, so I think I have found the path ((laughs))· that you meant. (incompr.) on the left side, there are gardens and stuff?*
...	...
R:	So, jetzt weiß ich auch, was du mit silbernen Brücken meinst. ((lacht))· *So now I know what you mean by "silver bridges". ((laughs))·*

Again several turns between S and R are needed to fix the CG which has been broken for some time. Similar to variant 1 roles between S and R are swapped: R is able to derive the location at which the misunderstanding occurred and derives the corrected route instruction ("after the bus stop I should have gone left, right?"). On arriving there,

R recognises several entities introduced earlier by S and, finally, establishes CG by explicitly referring to the "silver bridges" which have been introduced earlier by S.

While both variants include different reasons for non-established CGs, the solution for both variants is essentially the same: R and S swap roles, i.e. the receiver starts to introduce objects which are salient to her/him in the environment using EPCs and expects a confirmation by the sender in order to establish CG. Receivers will, then, continue to make route suggestions (see variant 2) or senders will continue to give instructions (see variant 1).

5.3 HCI Design Guidelines to foster CG establishing

Taken together, our corpus provides evidence for a general discourse pattern: S introduces a salient object by using an EPC and expects a confirmation by R about the recognition of this objects in the local environment. If this is the case, CG is established and S continues to give a route instruction. If, however, R cannot identify the salient object, it is not yet part of the CG content: Restoring CG is done by swapping roles between R and S, during which R primarily uses EPCs to refer to entities which either have already been part of the CG content or which are salient to her/him. Only when CG is re-established, S continues to provide route instructions.

Taken together, the dialogue analysis provide evidence that the discourse between sender and receiver is about continuously tracking the state of affairs with respect to explaining a route and localisation: Both, sender and receiver, can assume (A) or know (K) that the receiver's current position is correct (C) or wrong (W) in the context of the current navigation step. If and only if both, sender and receiver, know that the position is correct, the CG for the next instruction is successfully established. This means, a wayfinding assistance system acting as a sender has to implement a task model consisting of 16 unique states at any point of time in order to establish CG:

$$\left(\underbrace{\text{certainty quality}}_{\text{sender}} \underbrace{\text{certainty quality}}_{\text{receiver}} \right)$$

Given these definitions, each state can be represented as a quadruple (certainty = $\{A, K\}$ and quality = $\{C, W\}$). Thereby,

$$(K, C, K, C)$$

represents the desired final state[3] in which the CG is successfully established. In human-to-human dialogues, EPCs serve the pragmatic purpose of changing the current state of the grounding task to the final state. A wayfinding assistance system, which mimicks human discourse, must track the state of tasks for establishing CG continuously throughout the navigation process. It has to compute a belief (i.e. a probability distribution over all 16 states) each time it receives new input from users, regardless whether it is point-and-click or natural language based. Graphical user interfaces, however, have

---

[3]Despite its use in a different domain, a similar approach has been taken successfully in [47].

the advantage to avoid additional uncertainty introduced by misinterpretation of user utterances. In any case, a user interface should allow to update the task state and adhere to the workflow of human-to-human dialogues (see above).

From our point of view, it is, therefore, essential to use an easy to comprehend interface to establish CG in a sequential manner which mimicks human-to-human CG negotiations through EPCs. Specifically, we are able to provide two guidelines for the implementation of such a user interface. The use of salient objects is a prerequisite according to our and earlier evidence [see e.g. 15, 48, 49]. Selecting the most salient object in a given spatial context is, thereby, a challenge itself, because of personal preferences, personal salience, personality traits and spatial abilities [see e.g. 24].

Assuming that an object has been identified, the user interface should, first, present it and request explicit feedback from the user whether it is recognised. If the user fails to identify the object in the environment and gives appropriate feedback, this step should be repeated with other objects (backward-sorted with respect to their salience) up until the system can reliably assume that it agrees with the user on the current position. Clearly, system knowledge about the user position (e.g. based on GNSS) is insufficient to draw conclusions about the user's knowledge about and confidence in his/her current position. A route instruction can only be successful if both, the current position and the user knowledge about this position match. In these cases the route instruction can be based on an object match. In order to achieve this match, e.g. in a speech enabled interface, the system's output and instructions should be limited to the syntactic and semantic structure of EPCs because users are adapted to these expressions in navigation instructions as the dialogues indicate. Interrupting the voice output should be enabled (e.g. using a specific keyword) when providing feedback as the turns in our corpus show considerable overlap. A second guideline deals with the inversion of roles between sender and receiver. We provide evidence that CG is most commonly repaired by a receiver introducing new or referring to entities which are already part of the CG content. An ideal system, therefore, would be able to understand user input on salient objects, i.e. users can change the system's belief by providing environmental information. One potential way towards such system capabilities is to train the Natural Language Understanding module of a speech enabled navigation user interface to detect these pragmatic purposes and update the belief about the task state accordingly. It remains, however, an open research question how the system can match the information contained in user input with its knowledge about the environment: Comprehending user descriptions of their current environment as well as identification of the objects using the system's knowledge base are two unresolved challenges. A first step, however, may be the wayfinding assistance system's capability to understand that a user feels lost. One option to realize this capability in a GUI scenario is the implementation of an appropriate *emergency button* which initiates a dialogue to establish CG as outlined above: The system asks the user iteratively about surrounding objects (starting with the most salient one) until CG is established.

6 Discussion

In this work we present and analyse the POPRIS-corpus of turn-by-turn landmark-based route instructions and illustrate the importance of its linguistic properties for wayfinding assistance systems. The main contribution of this novel corpus is based on the full transcripts of the dialogues between pairs of participants who were unfamiliar with the experimental area. In addition to transcripts of utterances, the corpus contains sense of direction and personality data and is, therefore, expected to be valuable for a variety of disciplines.

The corpus data provides evidence for the dominance of EPCs to establish CG. Both, senders and receivers use these to introduce new salient objects or refer to those mentioned before in case CG is not established. Two variants were discussed in which CG is not established: In variant 1, the reason is a misunderstanding [see 50], i.e. the receiver confirms the recognition of different silver bridges than those the sender actually referred to. Variant 2, however, is a case of misinterpretation [see 50]: The sender refers to a right-hand bend not recognised by the receiver, whereas the sender cannot interpret the receiver's reference to the church.

Based on our empirical findings we derived a negotiation protocol between a user and an assistance system, which is suitable to establish CG. By explicitly requiring feedback on the recognition of a salient object before a route instruction is based on it, the proposed assistance system can help resolve ambiguities, avoid uncertainty as well as reduce cognitive load. Therefore, user feedback is requested at the decision point level and continues until CG is established. This is in contrast to [34], in which it is suggested to stop negotiations between system and user when the decision point level is reached. Generally speaking, it is interesting to see that the formulation and presentation of route instructions has seen major research interest; however, state-of-the-art user interfaces using these results are (with the exception of [34]) implemented on the assumption that the current position and the user knowledge about this position matches the system's assumptions/knowledge of both ((K, C, K, C), see above) and that this match is reached immediately each time a new instruction is presented. This is in sharp contrast to evidence, for example, [see 51] on the fact that maps generally show multiple possible interpretations, rendering errors in CG very likely. The evidence and approach we derive, however, is in-line with the guideline presented in [52], as the confirmation of CG resembles the confirmation of correct decision making. The closest approach to our proposed assistance system is proposed in [53]; Bauer describes an experiment during which participants were required to press a button labelled "Ziel erkannt" (goal perceived) [53, p. 89] when participants had the impression that they understood the landmark-based instruction as a whole. In this prototype a holistic view on an instruction is proposed, consequently. In contrast to this holistic view, however, we provide evidence for the fact that introducing a salient object using an EPC and providing an instruction should be two distinct steps. An assistance system, which mimics this way of human-to-human CG management, is likely to resolve trust issues, which might occur when a system uses references to landmarks (see e.g. [54] for this claim). In addition to that, since the negotiation phase requires to search for and recognise objects in the environment the acquisition of spatial knowledge is supported, thereby counterbalancing potential negative effects of extensive navigation system use [see e.g. 55].

The major limitation which applies in terms of implementing such systems is the availability of large scale salience values for objects – despite theoretical [see e.g. 21, 22, 24] and empirical [see e.g. 25, 56, 26, 23, 28] efforts. Currently, no large-scale geospatial database exists from which a wayfinding assistance system can extract candidate objects and use these for establishing the CG. The issue of how to fill a database with high quality entries is an open question for research and practice of navigation systems. VGI [57] approaches may possibly offer a way to acquire this kind of geospatial knowledge on a large scale.

7 Conclusion and Future Work

In this paper we shed light on the use of existential-presentative constructions to establish Common Ground (CG) in a wayfinding task and their implications for the design of wayfinding assistance systems. Specifically, we present a corpus of on-line route instructions given in turn-by-turn manner over the phone in-situ. Existential-presentative constructions (EPCs) appear to be the most common way of introducing salient objects (landmarks), based on which, first, CG is negotiated and, second, an actual route instruction is given referring to this object. This human-to-human behaviour is, from our point of view, a suitable means to base a system's behaviour on. A wayfinding assistance system should request explicit user feedback on whether the salient object it refers to is recognised or not. If not, it should continue to ask for recognition of other objects in order to establish CG. A route instruction should be issued only then. In addition to that, the user interface should allow users to initiate dialogues (e.g. by means of an *emergency button*) about objects they recognise in the environment in order to ensure correct self-positioning of the user. Three different strands of future work are planned. A first strand is dedicated to in-situ and virtual reality HCI studies in order to determine suitable interaction techniques for cases in which CG is and those in which it is not established. To this end, we plan a series of virtual reality and in-situ experiments. We, furthermore, plan to analyse how users of assistance systems make sense of route instructions based on think-aloud protocols and if/how this might differ from human-to-human route communication (in particular with respect to the information needs expressed using EPCs). These differences may have an impact on the efficiency of route instructions provided by the system. Finally, we intend to add an implementation of the task model for establishing the CG to an existing navigation system that currently cannot track a belief state based on user feedback: We plan to use the speech data we have collected to train a NLU component that helps in tracking the state of grounding tasks by analysing user utterances.

References

[1]   Reginald Golledge. "Human Wayfinding and Cognitive Maps". In: *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*. Ed. by Reginald G. Golledge. Baltimore, MA: The Johns Hopkins University Press, 1999, pp. 5–45.

[2]   Anna C. Farr et al. "Wayfinding: A simple concept, a complex process". In: *Transport Reviews* 32.6 (2012), pp. 715–743.

[3]   Ioannis Giannopoulos et al. "Wayfinding Decision Situations : A Conceptual Model and Evaluation". In: *Geographic Information Science. GIScience 2014. Lecture Notes in Computer Science*. Ed. by Matt Duckham et al. GIScience '14. Springer International Publishing, 2014, pp. 221–234. DOI: 10.1007/978-3-319-11593-1_15.

[4]   Thora Tenbrink. "Cognitive discourse analysis: Accessing cognitive representations and processes through language data". In: *Language and Cognition* 7.1 (2015), pp. 98–137.

[5]   Karl Rehrl et al. "An analysis of direction and motion concepts in verbal descriptions of route choices". In: *International Conference on Spatial Information Theory*. Springer. 2009, pp. 471–488.

[6]   Thora Tenbrink et al. "Communicative Success in Spatial Dialogue: The Impact of Functional Features and Dialogue Strategies". In: *Language and Speech* 60.2 (2017), pp. 318–329. DOI: 10.1177/0023830916651097.

[7]   David Crystal. *An Encyclopedic Dictionary of Language and Languages*. Oxford, 1994.

[8]   Kris Liu, Jean Fox Tree, and Marilyn Walker. "Coordinating Communication in the Wild: The Artwalk Dialogue Corpus of Pedestrian Navigation and Mobile Referential Communication". In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. Portorož, Slovenia: European Language Resources Association (ELRA), May 2016, pp. 3159–3166. URL: https://www.aclweb.org/anthology/L16-1504.

[9]   Jana Götze and Johan Boye. "SpaceRef: A corpus of street-level geographic descriptions". In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. Portorož, Slovenia: European Language Resources Association (ELRA), May 2016, pp. 3822–3827. URL: https://www.aclweb.org/anthology/L16-1605.

[10]   Anne H. Anderson et al. "The HCRC Map Task Corpus". In: *Language and Speech* 34 (1991), pp. 351–366.

[11]   Hamburger Zentrum für Sprachkorpora. *HAMATAC - the Hamburg MapTask Corpus*. 2010. URL: http://hdl.handle.net/11022/0000-0000-6330-A.

[12]   Simon Sauer and Anke Lüdeling. "Flexible Multi-Layer Spoken Dialogue Corpora. Special Issue: Compilation, Transcription, Markup and Annotation of Spoken Corpora". In: *International Journal of Corpus Linguistics* 21.3 (2016), pp. 419–438.

[13]   Jan Gorisch et al. "Aix Map Task corpus: The French multimodal corpus of task-oriented dialogue". In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Ed. by Nicoletta Calzolari

et al. Reykjavik, Iceland, 2015, pp. 2649–2652. ISBN: 978-2-9517408-8-4. URL: http://nbn-resolving.de/urn:nbn:de:bsz:mh39-39235.

[14] Kenny R. Coventry, Thora Tenbrink, and John Bateman, eds. *Spatial Language and Dialogue*. Vol. 3. Explorations in language and space. Oxford University Press, 2009. ISBN: 978-0-19-955420-1. DOI: 10.1093/acprof:oso/9780199554201. 001.0001. URL: https://doi.org/10.1093/acprof:oso/9780199554201.001.0001.

[15] Kristin L. Lovelace, Mary Hegarty, and Daniel R. Montello. "Elements of good route directions in familiar and unfamiliar environments". In: *Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science*. Ed. by Christian Freksa and David M. Mark. Lecture Notes in Computer Science. 1999, pp. 65–82. ISBN: 3-540-66365-7.

[16] Dieter Wunderlich. "How to get there from here". In: *Speech, place, and action* (1982), pp. 183–201.

[17] Wolfgang Klein. "Deixis and spatial orientation in route directions". In: *Spatial orientation*. Springer, 1983, pp. 283–311.

[18] Talmy Givón. "Syntax, vol. II". In: *Benjamins, Amsterdam* (2001).

[19] Jennifer D. Ewald. ""Do you know where X is?": Direction-giving and male/female direction-givers". In: *Journal of Pragmatics* 42.9 (2010), pp. 2549–2561.

[20] Susan Brennan, Katharina Schuhmann, and Karla Batres. "Entrainment on the move and in the lab: The walking around corpus". In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 35. 2013.

[21] Martin Raubal and Stephan Winter. "Enriching Wayfinding Instructions with Local Landmarks". In: *Geographic Information Science*. Ed. by Max Egenhofer and David Mark. Lecture Notes in Computer Science. Berlin / Heidelberg: Springer, 2002, pp. 243–259. ISBN: 978-3-540-44253-0.

[22] David Caduff and Sabine Timpf. "On the assessment of landmark salience for human navigation". In: *Cogn. Process.* 9.4 (2008), pp. 249–267.

[23] Markus Kattenbeck. "How Subdimensions of Salience Influence Each Other. Comparing Models Based on Empirical Data". In: *13th International Conference on Spatial Information Theory, COSIT 2017, September 4-8, 2017, L'Aquila, Italy*. Ed. by Eliseo Clementini et al. 2017, 10:1–10:13. DOI: 10.4230/LIPIcs. COSIT.2017.10. URL: https://doi.org/10.4230/LIPIcs.COSIT.2017.10.

[24] Eva Nuhn and Sabine Timpf. "A multidimensional model for selecting personalised landmarks". In: *Journal of Location Based Services* 11.3-4 (2017), pp. 153–180. DOI: 10.1080/17489725.2017.1401129.

[25] Birgit Elias. "Extraktion von Landmarken für die Navigation". PhD thesis. Hannover: Fakultät für Bauingenieurwesen und Geodäsie der Universität Hannover, 2006.

[26] Jana Götze and Johan Boye. "Learning landmark salience models from users' route instructions". In: *Journal of Location Based Services* 10.1 (2016), pp. 47–63. DOI: 10.1080/17489725.2016.1172739.

[27] Markus Kattenbeck. *Empirically Measuring Salience of Objects for Use in Pedestrian Navigation*. July 2016. URL: https://epub.uni-regensburg.de/34145/.

[28] Markus Kattenbeck, Eva Nuhn, and Sabine Timpf. "Is Salience Robustl A Heterogeneity Analysis of Survey Ratings". In: *10th International Conference on*

Geographic Information Science (GIScience 2018). Ed. by Stephan Winter, Amy Griffin, and Monika Sester. Vol. 114. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018, 7:1–7:16. ISBN: 978-3-95977-083-5. DOI: 10.4230/LIPIcs. GISCIENCE.2018.7. URL: http://drops.dagstuhl.de/opus/volltexte/2018/9335.

[29] Wolfgang Maass and D. Schmauks. "MOSES: an example for the modelling of spatial services by means of route description system". In: *Zeitschrift für Semiotik* 1.2 (1998), pp. 105–118. URL: http://iss.uni-saarland.de/de/publications.

[30] Wolfgang Maass. "How Spatial Information Connects Visual Perception and Natural Language Generation in Dynamic Environments: Towards a Computational Model". In: *A Theoretical Basis for GIS, Proc. of the Int. Conf. COSIT'95 Spatial Information Theory, Lecture Notes in Computer Science*. Ed. by A. Frank and W. Kuhn. 1995, pp. 223–240. URL: http://iss.uni-saarland.de/de/publications.

[31] Klaus-Peter Gapp and Wolfgang Maass. "Spatial Layout Identification and Incremental Descriptions". In: *Proc. of AAAI-94 Workshop on "Integration of Natural Language and Vision Processing", Seattle, WA*. Ed. by Paul Mc Kevitt. 1994, pp. 145–152. URL: http://iss.uni-saarland.de/de/publications.

[32] Martin Tomko and Stephan Winter. "Recursive construction of granular route directions". In: *Journal of Spatial Science* 51.1 (2006), pp. 101–115.

[33] Kai-Florian Richter and Alexander Klippel. "A Model for Context-Specific Route Directions". In: *Spatial Cognition IV. Reasoning, Action, Interaction*. Ed. by Christian Freksa et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 58–78.

[34] Kai-Florian Richter, Martin Tomko, and Stephan Winter. "A dialog-driven process of generating route directions". In: *Computers, Environment and Urban Systems* 32.3 (2008), pp. 233–245.

[35] Alexander Klippel, Kai-Florian Richter, and Stefan Hansen. "Cognitively ergonomic route directions". In: *Handbook of Research on Geoinformatics*. IGI Global, 2009, pp. 230–238.

[36] Alexander Klippel et al. "Urban granularities—a data structure for cognitively ergonomic route directions". In: *GEO* 13.2 (2009), pp. 223–247.

[37] Vivien Mast and Diedrich Wolter. "A probabilistic framework for object descriptions in indoor route instructions". In: *International Conference on Spatial Information Theory*. Springer. 2013, pp. 185–204.

[38] Paolo Fogliaroni et al. "Intersections of our world". In: *10th International Conference on Geographic Information Science*. Vol. 114. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. 2018, p. 3.

[39] Stefan Münzer and Christoph Hölscher. "Entwicklung und Validierung eines Fragebogens zu räumlichen Strategien". In: *Diagnostica* (2011).

[40] Daniel Danner et al. "Das Big Five Inventar 2: Validierung eines Persönlichkeitsinventars zur Erfassung von 5 Persönlichkeitsdomänen und 15 Facetten". In: *Diagnostica: Zeitschrift für psychologische Diagnostik und differentielle Psychologie* 65.3 (2019), pp. 121–132.

[41] W. Andrew Harrell, Jeffrey W. Bowlby, and Deana Hall-Hoffarth. "Directing Wayfinders With Maps: The Effects of Gender, Age, Route Complexity, and

Familiarity With the Environment". In: *The Journal of Social Psychology* 140.2 (2000), pp. 169–178. DOI: 10.1080/00224540009600456.

[42] Caroline Jones and Amit German. *Review of F4transkript, a simple interface for efficient annotation*. 2016. DOI: http://hdl.handle.net/10125/24701.

[43] Jochen Rehbein et al. *Handbuch für das computergestützte Transkribieren nach HIAT*. de. Hamburg: Universität Hamburg - Sonderforschungsbereich 538, 2004, p. 78. URL: http://nbn-resolving.de/urn:nbn:de:bsz:mh39-23681.

[44] Stefan Münzer, Benedict C.O.F. Fehringer, and Tim Kühl. "Validation of a 3-factor structure of spatial strategies and relations to possession and usage of navigational aids". In: *Journal of Environmental Psychology* 47 (2016), pp. 66–78. ISSN: 0272-4944. DOI: 10.1016/j.jenvp.2016.04.017.

[45] Caroline Féry and Manfred Krifka. *Information structure*. 2009.

[46] Manfred Krifka. "Basic notions of information structure". In: *Acta Linguistica Hungarica* 55.3-4 (2008), pp. 243–276.

[47] Sabrina Barko-Sherif, David Elsweiler, and Morgan Harvey. "Conversational agents for recipe recommendation". In: *Proceedings of the 5th ACM SIGIR Conference on Human Information Interaction and Retrieval, Vancouver, British Columbia, Canada on 14–18 March, 2020*. accepted for publication. 2020.

[48] Marie-Paule Daniel et al. "Testing the Value of Route Directions Through Navigational Performance". In: *Spatial Cognition & Computation* 3.4 (2003), pp. 269–289. DOI: 10.1207/s15427633scc0304_2.

[49] Ariane Tom and Michel Denis. "Language and Spatial Cognition: Comparing the Roles of Landmarks and Street Names in Route Instructions". In: *Applied Cognitive Psychology* 18 (2004), pp. 1213–1230. DOI: 10.1002/acp.1045.

[50] Susan McRoy. "Preface: Detecting, repairing and preventing human—machine miscommunication". In: *International Journal of Human-Computer Studies* 48.5 (1998), pp. 547–552.

[51] Christina Ohm, Manuel Müller, and Bernd Ludwig. "Evaluating indoor pedestrian navigation interfaces using mobile eye tracking". In: *Spatial Cognition & Computation* 17.1-2 (2017), pp. 89–120. DOI: 10.1080/13875868.2016.1219913. eprint: https://doi.org/10.1080/13875868.2016.1219913. URL: https://doi.org/10.1080/13875868.2016.1219913.

[52] Andrew J. May et al. "Pedestrian navigation aids: information requirements and design implications". In: *Personal and Ubiquitous Computing* 7.6 (2003), pp. 331–338.

[53] Christina Bauer. "Unterstützung der Orientierung im Innenbereich: Analyse landmarkenbasierter Karten-Interfaces anhand des Blickverhaltens der Nutzer". PhD thesis. University Regensburg, 2018.

[54] Tad T. Brunyé et al. "Where did it come from, where do you go? Direction sources influence navigation decisions during spatial uncertainty". In: *Quarterly Journal of Experimental Psychology* 68.3 (2015), pp. 585–607. DOI: 10.1080/17470218.2014.963131.

[55] Annina Brügger, Kai-Florian Richter, and Sara Irina Fabrikant. "How does navigation system behavior influence human behavior?" In: *Cognitive Research: Principles and Implications* 4.5 (2019). DOI: 10.1186/s41235-019-0156-5.

[56] Markus Kattenbeck. "Empirically Measuring Salience of Objects for Use in Pedestrian Navigation". In: *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*. SIGSPATIAL '15. Seattle, Washington: Association for Computing Machinery, 2015. ISBN: 9781450339674. DOI: 10.1145/2820783.2820820. URL: https://doi.org/10.1145/2820783.2820820.

[57] Michael F. Goodchild. "Citizens as sensors: the world of volunteered geography". In: *GeoJournal* 69.4 (2007), pp. 211–221.