# “Turn left” vs. “Walk towards the café”: When relative directions work better than landmarks

**Jana Götze and Johan Boye**

School of Computer Science and Communication, KTH Royal Institute of Technology, Sweden.

e-mail: jagoetze@kth.se, jboye@kth.se

**Abstract** An automatic mechanism that gives verbal navigation instructions to pedestrians in situ needs to take into account a number of factors. Besides giving the instruction at the right time and place, the information needs to be as unambiguous as possible for the user to both choose the correct path and be confident in doing so. Humans make extensive use of landmarks when describing the way to others and are more successful following instructions that include landmarks. We present a study comparing landmark-based instructions with relative direction instructions on pedestrians in a real city environment, measuring both objective and subjective success. We find that at some decision points, relative direction instructions work better. We present a method that uses openly available geographic data to predict which kind of instruction is preferable at a given decision point.

## 1 Introduction

Giving and following route instructions are tasks that most of us carry out regularly. Nowadays, many people have a smartphone with a built-in GPS receiver. This new state of affairs, along with the increasing coverage and quality of open geographic databases (like OpenStreetMap [11]), has opened up the possibility of implementing systems that give real-time verbal routing instructions for the city pedestrian. When there is no map at hand, receiving well-timed instructions at the scene can present an improvement compared to, say, having to look at printed instructions on a sheet of paper, or having to memorize the entire route beforehand. An advanced verbal routing system might also be interactive and able to understand requests for clarification from the user (e.g. [4, 13]). Such systems are in many cases able to provide an alternative explanation in case the user has not understood where to go next.

To minimize the number of misunderstood instructions and ensuing clarification requests, it is of great importance for system designers to come up with algorithms which will generate the most successful instruction in any given situation. By ‘successful’ we mean both objectively successful (in the sense of leading to as few navigational errors as possible) and subjectively successful (in the sense of user satisfaction and perceived confidence). Furthermore, instructions should be easy for the pedestrian to process. They should be interpretable without a map and, as we are aiming for practical applications, possible to generate from freely available geographic databases.

In this paper, we present a study carried out to provide some empirical basis for such an instruction-giving algorithm. Our aim is to compare the effectiveness of landmark-based instructions with relative direction (left/right/straight) instructions (and with combinations such as “Turn left towards the X”). Several studies have pointed out that landmarks play a special role in the communication of route directions [2, 5, 9, 17, 19, 21, 24]. However, direction givers do not always prefer to give a landmark-based instruction and by studying in situ route instructions given by people, we noted that this seems to depend on where the instruction is given. We are now interested in how such relative directions are interpreted by someone who has to follow them, and under which circumstances such instructions work well.

Furthermore, as a practical consideration, the most preferable (or salient) landmark might not be in the database (e.g. the billboard with the big clock is not currently (Nov. 2014) represented in OpenStreetMap (see Figure 1). Another possibility is that a salient landmark (e.g. a building) is present in the database, but the feature that distinguishes it (e.g. its red color) is not represented. Thus, there are reasons to explore other kinds of instructions than landmark-based ones.

**Fig. 1.** The billboard with the clock is a salient landmark not present in the geographic database.

# 2 Background

## *2.1 Route directions*

How people give and understand route directions has been studied extensively. While there is no single definition for good route directions, several aspects have been found to play a role:

A route is typically split into several segments that are then verbalized [6]. These verbalized directions can be instructions to take a particular action, such as “walk” or “turn”, or descriptions of the environment like “There is a red building to your left.” that help the direction follower to identify where an action is to be carried out or whether he is still on the right way [2, 9]. The order of the directions should reflect the linear order in which the route is traversed [3]. These conclusions have been drawn from verbal or written route directions that have been produced prior to someone following them, i.e. the route follower had to memorize the directions before carrying them out.

Aspects of route direction following have also been studied in a setting where directions are given while the follower proceeds along the route. Hund & Minarik [12] as well as May & Ross [16] have studied route directions for car navigation, and several systems have been built for direction giving in a virtual environment (cf. Striegnitz et al. [22]), focusing mostly on the generation of referring expressions. For directions given to pedestrians, studies typically involve the direction follower to find a route on a map or draw a route on paper [14, 17, 24].

## *2.2 The role of landmarks in route directions*

That landmarks play a special role in the communication of route directions has been demonstrated in several studies. Landmarks are a means to identify crucial points along the route where turning actions need to be taken or could be taken (decision points), as well as to locate the beginning and the end of the route [17]. Landmarks also play a role in the descriptive part of route directions, to locate other landmarks, and to confirm that the follower has correctly executed a turn [8]. Salient landmarks can also be used as a criterion to find the best route in terms of how easy it will be to follow [5, 21].

The term landmark is usually used for a visually salient object along the route that is easily identifiable by some outstanding feature, such as its size, color, or special function. While street names do not serve as good landmarks because they cannot easily be identified [24], streets themselves can serve as landmarks if they are described vividly in terms of their outstanding features [25].

Raubal and Winter [19] try to compute the salience of landmarks using different salience measures to reflect visual, structural, and semantic salience. However, the information they use is often not readily available, such as a landmark's color, façade, or cultural importance. We [10] have used information that is freely available from OpenStreetMap, such as a landmark's position and type, to compute its salience.

### *2.3 Comparing different types of instructions*

While it is undisputed that landmarks play a vital role in making route directions easy to follow, automatically choosing an appropriate landmark in a given situation is not an easy task. There is often insufficient information available to decide which landmark stands out in terms of salience, such as its color, height, or special function.

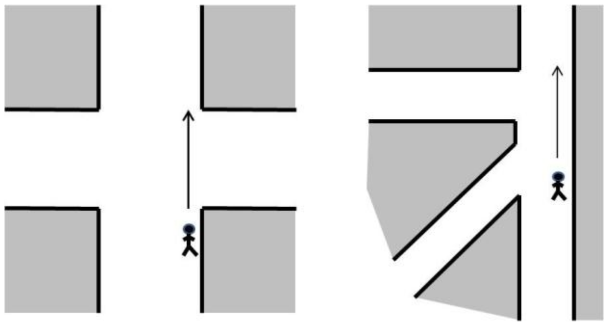
Looking at human route directions, we find that at some decision points, the describers seem to prefer to not use a landmark. In Daniel & Denis' study [24], about 14-20% of route directions did not include a landmark. In Rehrl's study [20], in which participants were giving directions in situ, while walking along the route, only 1-7.4% of the directions were instructions without a landmark. However, this reflects only the preference of the direction giver to include landmarks. We are interested in knowing how useful such an instruction is for a direction follower.

## 3 Modeling the spatial environment

Let us consider the natural hypothesis that if the decision point has a simple configuration, such as a T-intersection or a four-way intersection where all streets meet at right angles, an instruction containing a relative direction will be sufficient. By contrast, if the decision point has a more complex configuration, such as many streets meeting at non-right angles, the inclusion of a landmark in the instruction will yield higher confidence and fewer errors. For instance, in the situation depicted in Figure 2A below, the instruction "please turn left" would be sufficient, whereas in Figure 2B it would not.

Though it is easy to draw such idealized maps and make conjectures about which instructions would be suitable in the clear-cut cases, it is less obvious how to devise a general algorithm for deciding, for any given situation, whether an instruction should be couched as a relative instruction or a landmark-based instruction. Moreover, such an algorithm must work in real time, and only use the information present in open geographic databases like OpenStreetMap.

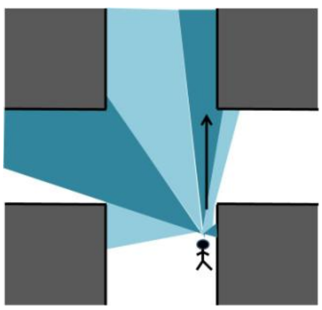
**Fig. 2A. (left)** A situation where "turn left" is probably clearly comprehensible, and **2B (right)** where it is not. The arrows indicate the walking direction.



OpenStreetMap has (basically) two data types: *nodes* and *ways*. Ways are sequences of nodes, used for representing a wide variety of objects, such as roads, squares, areas and buildings (in the three latter cases, the first node in the sequence is the same as the last node, and hence the way forms the perimeter of a polygon). An intersection between two streets is represented by the point where the ways corresponding to the streets meet. However, the situation in Figure 2B (and similar situations such as roundabouts), that for the human eye constitutes a single complex intersection, have no explicit representation in OpenStreetMap. This has to be taken into account when designing an algorithm for instruction generation from geographical databases.

Our system for pedestrian routing instructions [10] contains a visibility engine, which can quickly perform line-of-sight computations. We use this engine to compute, for a given situation, **how far the closest building is** in every direction from -100 degrees to 100 degrees, relative to the user's current direction (cf. Figure 3). This vector of 201 numbers thus constitutes a crude representation of the user's spatial surroundings. One of our aims is now to investigate whether there is a systematic relationship, on the one hand between the numbers of this vector for a given situation, and on the other hand how test subjects perceive a relative instruction in that situation.

**Fig. 3.** The considered sectors of the user's field of vision



In order to reduce the dimensionality of the input data, we divide the user's field of vision into 7 subsectors: straight ahead (-10 to 10 degrees), as well as 11-40, 41-70, and 71-100 degrees on either side, and only consider the maximal distance in each subsector. For example, a typical such vector could be (252, 20, 10, 200, 14, 2, 1) from left to right, meaning that the user has a free line of sight extending 252 metres to the left in the sector -100 to -71 degrees, 20 metres in the sector -70 to -41 degrees, and so on.

# 4 Experiment

In order to gain an insight into where it is preferable to use a landmark in an instruction, we asked a number of subjects to follow different kinds of route instructions. After each instruction they received, they rated their confidence in knowing which direction to choose. The instructions they received consisted of either a relative direction ("Turn left/right/straight"), or a landmark ("Walk towards the school"), or they combined both pieces of information.

**Materials**

We have determined two routes with slightly different street layout for this study. Route I contains 14 decision points with more complex configurations, such as a roundabout or several streets turning into similar directions. Route II contains 13 decision points with simpler configurations, where streets meet at right angles.

We constructed two sets of instructions for both routes. The first set contained only relative instructions of the form "Turn/Go <direction>". We call this set the *direction set D*. The second set, the *landmark set L*, contained instructions that used only a landmark, e.g. "Walk towards the school". For Route I, we also constructed a third set that combined both pieces of information into one, we call this set the *landmark+direction set LD*. The instructions for each route were then randomly arranged into two versions (Route II) of route instructions (three versions for Route I) with the only restriction that there should not be more than three instructions of the same kind following each other. Table 1 shows example instructions for each of the three sets of instructions.

In this study, we are focussing on only one part of route directions, the action prescriptions (cf. [8]). For route directions to be complete, they need to also contain a description of how to identify the decision point where an action is to take place. Here, we are only interested in the use of landmarks to carry out the action at a decision point. In the navigation system described below, we have access to the pedestrian's position through the GPS signal and can take part of the burden to locate the decision point off the pedestrian by saying "walk until I say stop".

**Table 1.** Example instructions for each of the three instruction types.

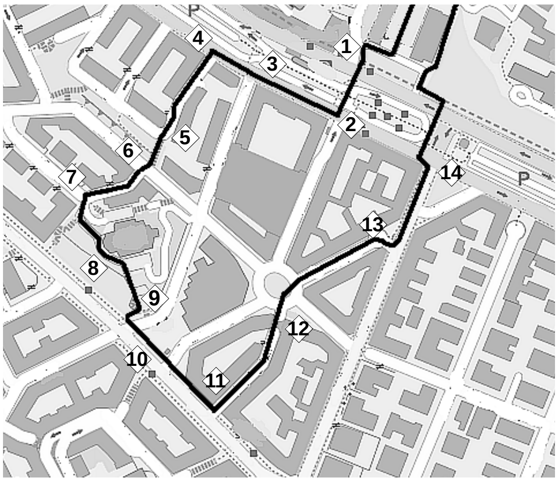
| Type | Example Instructions |
|---|---|
| Direction | Go straight.<br>Turn right. |
| Landmark | Go towards the school.<br>Go down the stairs.<br>Walk along the big street. |
| Direction+Landmark | Go straight towards the school.<br>Turn left towards the bicycles.<br>Go down the stairs on the right. |

Route I can be seen in Figure 4. This route, as well as the landmarks for the two landmark-based conditions, comes from one of our previous experiments in which participants were asked to describe the route while they were walking along it [1]. Route I was about 1.1 km long and contained 4 left turns, 3 right turns and 7 continuations, i.e. decision points where no change of direction occurred. Decision points were included at each point where a change of direction could occur. Most of the continuations were explicitly mentioned by the participants of our previous experiment and therefore included as decision points. This segmentation also conforms to the way that an automatic system could choose decision points. In [10] all nodes that initially belong to the route are chunked based on visibility and a limit on distance between them. If the angle between two segments is very small, the direction will be “straight”.

Route II contains 5 left turns, 6 right turns, and 2 continuations. For this route, continuations were omitted as explicit instructions, instead, the participant received a confirmation of the form “You are doing fine. Please keep walking straight”. This confirmation was not rated by the participants. The landmarks are selected manually and vary between types and names of shops, and descriptions of streets.

The participants were equipped with a Samsung Galaxy S4 mobile phone running an application that sent the participant’s position and speech data and received text that was synthesized by the phone’s Text-To-Speech application (TTS). The experimenter took the role of a wizard in a so-called Wizard-of-Oz (WoZ) data collection [7], and interacted with the participant by selecting what to say from a pre-determined set of utterances. When the participant reached a decision point, the experimenter selected the corresponding button that started a small dialog with the participant, asking him to stop, listen to the instruction, and rate it. An additional window allowed the experimenter to monitor the participant’s position trail as well as possibly problematic situations such as interruptions in the connection between the phone application and the WoZ interface.

**Fig. 4.** A map of Route I, containing 14 decision points. Map source: © OpenStreetMap contributors




## Participants

18 participants (6 female and 12 male, average age=29, SD=3.3) followed Route I in return for a cinema ticket. 14 participants (6 female, 8 male, average age=24.9, SD=2.3) followed Route II in return for course credit. For both routes, participants were randomly assigned one of three (Route I) or two (Route II) different sets of route instructions. None of them had participated in any of our previous experiments. None of them followed both routes.

## Procedure

Participants were asked to follow a set of instructions, each of which described the next direction to take. Each instruction was rated before they carried out the action it described.

The participants, after receiving an instruction, were asked to choose the direction that they thought was correct and walk straight until the system told them to stop to get the next instruction. In this way we could make sure that everyone got the instruction at the same point and that the spatial environment at the time of the instruction was stable and did not change due to the participant's movement[1]. Until the participant had rated the instruction, he or she was standing still at a decision point.

[1] Note that there is still some variability in the GPS signal.

The process of the experiment was the following: Each participant was equipped with a mobile phone running an application that connected to software running on the experimenter's computer, as well as a headset. The participant was informed that she would receive a set of navigation instructions through the phone, which she would have to rate on a scale from 1 to 5, reflecting her confidence in knowing which action to perform. 1 corresponds to the lowest possible confidence ("not confident at all"), 5 to the highest ("very confident"). The participant was then asked to carry out the action and was welcomed to leave spoken comments. She was asked to keep walking straight until the system asked her to stop for the next instruction.

Each participant was asked to step outside the building, where the experiment started with the first instruction. The first three or four instructions[2] were training instructions to accustom the participant to the synthetic voice and the order in which things were happening. They were not informed about the actual start of the experiment. At each decision point, the participant received detailed task instructions. Some of these were very detailed and as soon as the participant signalled that she was comfortable with the task, e.g. by barging in, she received a shorter version of the task description.

**Collected data**

Of the 18 participants, 14 completed all 14 route segments. This results in a total of 232 confidence ratings. A rating is annotated with an error tag if the participant chose another than the intended next route segment. Furthermore, we have time-stamped log files that contain GPS data, as well as all instructions and other prompts (e.g. those that ask for a rating) played to the participant.

## 5 Experiment results

The upper part of Table 2 shows an overview of the confidence ratings of Route I that the users gave for each of the 3 conditions, Landmark (L), Direction (D), and Landmark+Direction (LD), sorted by the kind of turn the follower had to take (*straight*: ↑, *left*: ←, and *right*: →). Average ratings for a segment in a given condition consist in 2 cases of 4 ratings, in 16 cases of 5 ratings, and in 24 cases of 6 ratings, and range from 2.17 to 5.0. Individual ratings range from 1 to 5. The average rating for all segments and conditions is 4.14, the median rating is 5.
The first seven rows show the continuation segments in which no change of direction occurred. In five of these cases, the D instruction was rated highest, with av-

2 This varied due to some construction work that we needed to circumnavigate to get the participants to the starting point.

erage ratings ranging from 4.2 to 5.0. In the other two continuation segments, the D instruction was rated 4.6 and 4.83, i.e. received a high rating as well. The overall average in ratings is 4.72. L instructions received an average of 3.85, this difference is significant ($t(61) = 3.53$, $p<.001$). The average rating for the LD instructions in these segments is 4.29. Instructions that included a landmark additionally to the relative direction were rated lower ($t(67) = 2.19$, $p<.04$). The difference in ratings between the L and the LD condition is not significant ($t(72) = 1.62$, $p=.11$).

In the segments where a change of direction occurred (cf. the middle of Table 2), the D instructions are rated highest only once, in segment 11 (4.33). The LD instructions receive higher ratings (4.44) than both the D instructions ($t(77) = 2.59$, $p<.02$) and L instructions ($t(68) = 2.19$, $p<.04$).

These two kinds of segments (direction change vs. no direction change) also differ in the number of errors that were made. Followers made more errors in the segments with a direction change, with 8 errors occurring after a D instruction, 6 occurring after an L instruction and 2 after an LD instruction. Errors in the segments without a direction change were never made in the D condition. The overall ratings for all segments differ between the LD instructions (4.37) and the L instruction (3.85; $t(142) = 2.73$, $p<.01$).

The lower part of Table 2 shows the results for Route II. We have collected 73 ratings for the D condition and 75 ratings for the L condition (14 participants have followed the instructions). Average ratings for a segment in a given condition consist in 2 cases of 5 ratings and in all other cases of 7 ratings. Average ratings range from 2.57 to 5.0, individual ratings range from 1 to 5. The overall average of ratings is 4.40, the median rating is 5. The landmark-based instructions receive a confidence rating of 3.92 and the relative directions an average of 4.89 ($t(89) = 5.96$, $p<.0001$). Only three errors have been made, all after a landmark-based instruction.

The confidence ratings and errors that the participants following Route I made suggest that a relative direction is the best choice when going straight while followers are more confident with a landmark instruction if there is a change of direction. Overall, including both kinds of information works best both in terms of confidence ratings and number of errors.

The fact that a simple "straight" works best when no change of direction occurs is not surprising, especially for this particular route. Most decision points are complex intersections and going straight is both the simplest alternative to identify at many decision points and the default action, even without any instruction. An example for this can be found in segment 5, when the relative direction receives a lower average rating than the landmark-based instruction. The path that the participants are asked to follow consists of a flight of stairs that clearly lead away from the road they are walking along and that are easy to identify as "stairs" (hence the high L rating of 5.0). On the other hand, this is also the only possibility to continue straight and the D rating is not much lower (4.6).

**Table 2.** Route I average ratings at each of the decision points for each of the three strategies (D: relative direction, L: landmark, LD: direction and landmark combined)

| | | | Confidence Ratings | | | | |
|---|---|---|---|---|---|---|---|
| | Turning Direction | Segment | Direction (D) | Landmark (L) | Land-mark+Direction (LD) | Average[3] | # errors |
| Route I | ↑ | 1 | 5.00 | 3.67 | 3.83 | 4.12 | 1 |
| | ↑ | 3 | 4.40 | 2.33 | 3.60 | 3.38 | 1 |
| | ↑ | 5 | 4.60 | 5.00 | 4.83 | 4.82 | |
| | ↑ | 6 | 4.83 | 4.50 | 5.00 | 4.76 | 1 |
| | ↑ | 9 | 4.20 | 2.17 | 3.33 | 3.18 | 2 |
| | ↑ | 10 | 5.00 | 4.67 | 4.80 | 4.82 | |
| | ↑ | 14 | 5.00 | 4.80 | 4.80 | 4.86 | 1 |
| | ↑ Average | | 4.72 | 3.85 | 4.29 | 4.27 | |
| | ← | 4 | 4.60 | 4.83 | 5.00 | 4.82 | 1 |
| | ← | 7 | 2.80 | 3.83 | 3.83 | 3.53 | 3 |
| | ← | 11 | 4.33 | 4.20 | 4.00 | 4.18 | 2 |
| | ← | 13 | 3.50 | 2.50 | 4.00 | 3.27 | 3 |
| | → | 2 | 4.67 | 2.40 | 4.83 | 4.06 | 2 |
| | → | 8 | 3.17 | 5.00 | 4.67 | 4.24 | 2 |
| | → | 12 | 3.00 | 4.00 | 4.67 | 3.88 | 3 |
| | ←/→ Average | | 3.74 | 3.82 | 4.44 | 4.01 | |
| | | all | 4.22 | 3.85 | 4.37 | 4.14 | 22 |
| Route II | ← | 2 | 4.71 | 4.71 | | 4.71 | |
| | ← | 6 | 5.00 | 3.86 | | 4.27 | |
| | ← | 9 | 5.00 | 4.00 | | 4.50 | 1 |
| | ← | 12 | 4.86 | 2.57 | | 3.71 | 1 |
| | ← | 13 | 5.00 | 4.86 | | 4.93 | |
| | → | 3 | 5.00 | 3.86 | | 4.43 | |
| | → | 5 | 4.40 | 3.86 | | 4.08 | |
| | → | 7 | 5.00 | 3.60 | | 4.42 | |
| | → | 8 | 4.86 | 3.71 | | 4.29 | |
| | → | 10 | 5.00 | 4.71 | | 4.86 | |
| | → | 14 | 4.86 | 3.29 | | 4.07 | 1 |
| | ←/→ Average (all) | | 4.89 | 3.92 | | 4.40 | 3 |

However, this pattern may result from this particular route: most decision points are not simple four-way or T-intersections but more complex configura-

[3] Recall that in some cases, the number of ratings for each strategy differs in the same segment.

tions, like a roundabout where six streets meet and thus a relative direction is ambiguous, even if it contains an additional modifier, such as “slightly right”. Using a relative direction might work considerably better at a simple decision point, where streets meet at right angles. At the same time, some of the landmarks in the continuation segments may be inadequate and cause low confidence ratings because the follower cannot unambiguously identify them. We picked the landmarks from other peoples’ descriptions, but did not account for whether they included the direction explicitly as well.

Route II contains decision points with simpler configurations where all streets meet at right angles. The ratings that we have collected suggest that for this route, the relative directions result in higher confidence than the landmark-based instructions.

## 6 Deciding what instruction to give

The confidence ratings we obtain from these studies suggest that both the street configuration at the decision point and the kind of action the pedestrian is supposed to carry out (turning left or right) influence their confidence in a certain type of instruction. A method that can predict the confidence score for a relative direction instruction from the geographical context and the turning direction will be a useful tool to generate the most helpful navigation instruction in a given situation. We show how we can employ linear regression to build such a model from the data we have collected.

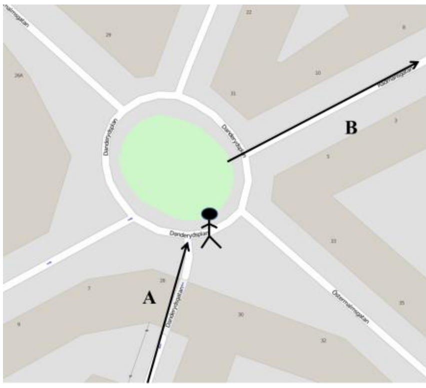
The dependent (output) variable we want to be able to predict is the confidence score for a relative direction (D) instruction. As explained previously, this score is a number 1-5 representing how intelligible a relative instruction is perceived in the particular situation at hand. The independent (input) variables should reflect two aspects of the decision point: the street configuration from the follower’s point of view, and the turning direction (we will only consider turns here and disregard the situations where the user is supposed to walk straight). We have shown in section 3 how we can model the geographical context: Each decision point is represented by a vector of seven numbers, indicating the farthest distance to a building (or similar three-dimensional structure) in different directions from the user’s current bearing. To include information about the turning direction into this vector, we are rearranging the vector as follows. The first three numbers represent the distances in the turning direction, the fourth number represents the distance straight ahead, and the last three numbers represent the distances on the side opposite of the turning direction. We thus have 7 independent variables $x_1 \ldots x_7$, representing distances in the various directions, and one dependent variable $y$, representing the estimated confidence score. In order not to over-fit the model, we use the integer value of the natural logarithm of the distances rather than the exact values of the distances (i.e. 26.84 metres would be represented by the value 3).

From our data, we have 18 route segments in which a change of direction occurred. We thus have 18 data points from which to, using linear regression, fit a linear function able to predict the average confidence scores that were given for a segment. To avoid over-fitting we are using 3-fold cross-validation (i.e. a model is trained using 2/3 of the data, and evaluated on the remaining 1/3; this process is repeated 3 times). The three models all show a high correlation between the seven independent variables (the line of sight) and the dependent variable (the rating); for the overall model (using all the data) we obtain a correlation coefficient of 0.91. The mean absolute error is 0.73. The overall model we obtain for all data points is the following equation:

$$y = 4.82 - 0.16x_1 - 0.16x_2 + 0.07x_3 + 0.24x_4 - 0.16x_5 - 0.04x_6 + 0.01x_7$$

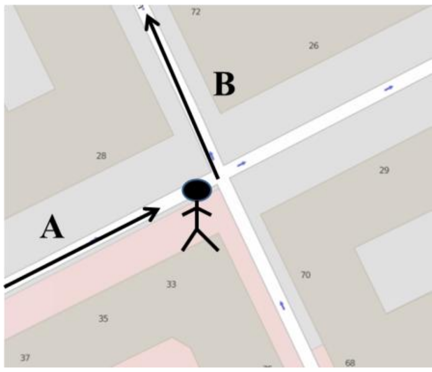
In the example in Figure 5, a route segment from Route I, arrow A indicates the follower's current bearing. This arrow corresponds to 0 degrees in our representation of the geographical context. The next route segment is reached via a right turn, but in this roundabout situation, other right turns are possible. In our experiment, the participants have given an average rating of 3.0 for the instruction "turn slightly right". The input vector for this particular situation is (8,5,8,5,7,7,8), and our regression model predicts a score of 3.18. Note that the average confidence scores do not span the whole range between 1 and 5. The lowest average for a segment is 2.8, the highest is 5.0. A score of 3 is thus a low score, meaning that in this route segment, a relative direction is not to be preferred.

**Fig. 5.** An example route segment from Route I




Consider by contrast the situation depicted in Figure 6. Here, participants gave an average rating of 5.0 for the relative instruction, which is not surprising given the simple street configuration. The input vector for this particular situation is (5,4,8,8,5,3,8), and the model predicts a value of 5.02. Here, an automatic system could confidently generate an instruction like "Turn left here".

**Fig. 6.** An example route segment from Route II




## 7 Conclusion

Our data collection shows that different kinds of instructions do not work equally well at all decision points. In particular, instructions that avoid a landmark and use only a relative direction like “left” or “right”, seem to be preferred at some decision points, particularly those with a simple configuration where streets meet at right angles.

We are suggesting a linear regression model to predict automatically whether such an instruction will be easy for a follower to interpret. The information we are using for this prediction is easily obtainable, e.g. from a database like OpenStreetMap, making it feasible to compute the confidence measure in real time. It is much easier to calculate the relative direction from a map than deciding which landmark to refer to and we are planning to include this automatic computation into our direction-giving system. If the predicted confidence score is at least 4, the system can be reasonably sure that the relative direction will be interpreted in the desired way.

## References


[1] Albore, A., Boye, J., Fredriksson, M., Götze, J., Gustafson, J. & Königsmann, J. Final pedestrian behaviour component. Project deliverable, Spacebook EU 7th framework project 270019, 2013.

[2] Allen, G. L. From knowledge to words to wayfinding: Issues in the production and comprehension of route directions. *Spatial Information Theory. A Theoretical Basis for GIS*, 363–372, 1997.

[3] Allen, G. L. Principles and practices for communicating route knowledge. *Applied Cognitive Psychology*, 2000, *14*, 333–359.

[4] Boye, J., Fredriksson, M., Götze, J., Gustafson, J. & Königsmann, J. Walk This Way: Spatial Grounding for City Exploration. *Natural Interaction with Robots, Knowbots and Smartphones, Springer New York,* 2014, 59–67.

[5] Caduff, D. & Timpf, S. The Landmark Spider: Representing Landmark Knowledge for Wayfinding Tasks. *AAAI'05 Spring Symposium*, 30–35, 2005.

[6] Couclelis, H. Portugali, J. (Ed.) Verbal Directions for Way-Finding: Space, Cognition, and Language. *The Construction of Cognitive Maps*, 1996, *32*, 133–153.

[7] Dahlbäck, N. & Jönsson, A. Empirical studies of discourse representations for natural language interfaces. *Proceedings of European Chapter of the Association for Computational Linguistics*, 1989, 291–298.

[8] Daniel, M.-P. & Denis, M. Spatial descriptions as navigational aids: a cognitive analysis of route directions. *Kognitionswissenschaft*, 1998, 7, 45–52.

[9] Denis, M. The description of routes: A cognitive approach to the production of spatial discourse. *Current Psychology of Cognition*, 1997, 16, 409–458.

[10] Götze, J.& Boye, J. Deriving Salience Models from Human Route Directions. *Workshop on Computational Models of Spatial Language Interpretation and Generation (CoSLI-3)*, 2013, 36–41.

[11] Haklay, M. & Weber, P. OpenStreetMap: User-Generated Street Maps. *Pervasive Computing, IEEE,* 2008*, 7*, 12-18.

[12] Hund, A. M. & Minarik, J. L. Getting From Here to There: Spatial Anxiety, Wayfinding Strategies, Direction Type, and Wayfinding Efficiency. *Spatial Cognition & Computation*, 2006, *6*, 179–201.

[13] Janarthanam, S., Lemon, O., Liu, X., Bartie, P., Mackaness, W., Dalmas, T. & Götze, J., Integrating location, visibility, and Question-Answering in a spoken dialogue system for Pedestrian City Exploration. In *Proceedings of SIGDIAL*, 2012, South Korea.

[14] Li, R., Fuest, S. & Schwering, A. The effects of different verbal route instructions on spatial orientation. *17th AGILE Conference on Geographic Information Science,* 2014.

[15] Lovelace, K., Hegarty, M. & Montello. Elements of Good Route Directions in Familiar and Unfamiliar Environments. Spatial Information Theory. *Cognitive and Computational Foundations of Geographic Information Science*, 1999, *1661*, 65–82.

[16] May, A. J. & Ross, T. Presence and Quality of Navigational Landmarks: Effect on Driver Performance and Implications for Design. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 2006*, 48*, 346–361.

[17] Michon, P.-E. & Denis, M. When and Why Are Visual Landmarks Used in Giving Directions? *Spatial Information Theory,* 2001*, 2205*, 292–305.

[18] Padgitt, A. J. & Hund, A. M. How good are these directions? Determining direction quality and wayfinding efficiency. *Journal of Environmental Psychology* , 2012*, 32*, 164–172.

[19] Raubal, M. & Winter, S. Enriching Wayfinding Instructions with Local Landmarks. *Geographic Information Science*, 2002, 2478, 243–259.

[20] Rehrl, K., Leitinger, S., Gartner, G. & Ortag, F. An Analysis of Direction and Motion Concepts in Verbal Descriptions of Route Choices. *Spatial Information Theory*, 2009, 5756, 471–488.

[21] K.-F. Richter. Context-Specific Route Directions — Generation of Cognitively Motivated Wayfinding Instructions. DisKi 314 / SFB/TR 8 Monographs Volume 3, 2008.
[22] Striegnitz, K., Denis, A., Gargett, A., Garoufi, K., Koller, A. & Theune, M. Report on the Second Second Challenge on Generating Instructions in Virtual Environments (GIVE-2.5). *13th European Workshop on Natural Language Generation*, 2011, 270–279.
[23] Tenbrink, T., Ross, R. J., Thomas, K. E., Dethlefs, N. & Andonova, E. Route instructions in map-based human–human and human–computer dialogue: A comparative analysis. *Journal of Visual Languages & Computing*, 2010, 21, 292–309.
[24] Tom, A. & Denis, M. Language and spatial cognition: comparing the roles of landmarks and street names in route instructions. *Applied Cognitive Psychology*, 2004, 18, 1213–1230.
[25] Tom, A. C. & Tversky, B. Remembering Routes: Streets and Landmarks. *Applied Cognitive Psychology,* 2012*, 26*, 182–193.