Comparing Spoken Language Route Instructions for Robots across Environment Representations

Matthew Marge
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
mrmarge@cs.cmu.edu

Alexander I. Rudnicky
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
air@cs.cmu.edu

Abstract

Spoken language interaction between humans and robots in natural environments will necessarily involve communication about space and distance. The current study examines people's close-range route instructions for robots and how the presentation format (schematic, virtual or natural) and the complexity of the route affect the content of instructions. We find that people have a general preference for providing metric-based instructions. At the same time, presentation format appears to have less impact on the formulation of these instructions. We conclude that understanding of spatial language requires handling both landmark-based and metric-based expressions.

1 Introduction

Spoken language interaction between humans and robots in natural environments will necessarily involve communication about space and distance. It is consequently useful to understand the nature of the language that humans would use for this purpose. In the present study we examine this question in the context of formulating route instructions given to robots. For practical purposes, we are also interested in understanding how presentation format affects such language. Instructions given in a physical space might differ from those given in a virtual world, which in turn may differ from those given when only a schematic representation (e.g., a map or drawing) is available.

There is general agreement that landmarks play an important role in spatial language (Daniel and Denis, 2004; Klippel and Winter, 2005; Lovelace et al., 1999; MacMahon, 2007; Michon and Denis, 2001; Nothegger et al., 2004; Raubal and Winter, 2002; Weissensteiner and Winter, 2004). However, landmarks might not necessarily be used uniformly in instructions across presentation formats. For example, people may use objects in the environment as landmarks more often when they do not have a good sense of distance in the environment. Behaviors related to spatial language may change based on the complexity of the route that a robot must take. This could be due to a combination of factors, including ease of use and personal assessment of a robot's ability to interpret specific distances over landmarks.

Several studies have investigated written or typed spatial language (e.g., MacMahon et al., 2006; Koulori and Lauria, 2009; Kollar et al., 2010). In addition, Ross (2008) studied models of spoken language interpretation in schematic views of areas. In the current study we focus on close-range spoken language route instructions.

2 Related Work

Interpreting spatial language is an important capability for systems (e.g., mobile robots) that share space with people. Human-human communication of spatial language has been extensively studied. Talmy (1983) proposed that the nature of language places constraints on how people communicate about space with others (i.e., *schematization*). Spatial descriptions are primarily influenced by how reference objects fit along fundamental axes that exhibit clear relationships with the target, and secondly by the salience of references (Carlson and Hill, 2008). People also tend to keep their spatial descriptions consistent after making an initial choice of strategy based on any existing relationships between the target to be described and other references (Vorwerg, 2009).

Proceedings of SIGDIAL 2010: the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue, pages 157–164,
The University of Tokyo, September 24-25, 2010.

(a) (b) (c)

Figure 1. Stimuli from the (a) schematic, (b) virtual, and (c) real-world scene experiments. Each scenario has 2 robots, Mok (left) and Aki (right). Mok is the actor in all scenarios. Outlined are possible destinations for Mok.

Studies involving spatial language with robots have thus far focused on scenarios where one robot is moved around an area using spatial prepositions (Stopp et al., 1994; Moratz et al., 2003) and further with landmarks (Skubic et al., 2002; Perzanowski et al., 2003). A number of these approaches, however, were crafted by the designers of the robots themselves and not necessarily based on an understanding of what comes naturally to people. Indeed, Shi and Tenbrink (2009) found that a person's internal linguistic representations may differ significantly from what a robot is capable of interpreting. Bugmann et al. (2004) motivated the concept of *corpus-based robotics*, where spontaneous spoken commands are collected and in turn used for designing the functionality of robots. They collected natural language instructions from people commanding robots in a miniature of a real-world environment. Our approach follows this same reasoning; we explore naturally occurring spatial language through route instructions to robots in three distinct formats (schematic, virtual, and natural environments).

3 Method

We designed and conducted three experiments using a navigation task that required the participant to "tell" a robot how to move to a target location. We varied the presentation formats of the stimuli (two-dimensional schematics, three-dimensional virtual scenes, real-world areas in-person). In each variant, the participant observed a static scene depicting two robots ("Mok" and "Aki") and a destination marker. The participant's task was to move Mok to the target destination using spoken instructions. Participants were told to act as if they were an observer of the scene but that were themselves not present in the scene; put otherwise, the robots could hear participants but not see them (and thus the participant could not figure in the instructions).

The experiment instructions directed participants to assume that Mok would understand natural language and were told to use natural expressions to specify instructions (that is, there was no "special language" necessary). Participants were told that they could take the orientations of the robots into account when they formulated their instructions. They were moreover asked to include all necessary steps in a single utterance (i.e., a turn composed of one or more spatial language commands). The robots did not move in the experiments.

Since our aim was to learn about spoken language route instructions, all participants recorded their requests using a simple recorder interface that could be activated while viewing the scene. A standard headset microphone was used. To avoid self-correction while speaking, the instructions directed participants to think about their instructions before recording. Participants could playback their instructions, and re-record them if they deemed them unsatisfactory. All interface activity was time-stamped and logged.

3.1 General variations

In their work, Hayward and Tarr (1995) found that people used spatial language with reference to landmarks most often and found it most suitable when the objects in a scene were horizontally or vertically aligned. We systematically varied three elements of the stimuli in this study: the orientations of the two robots, Mok and Aki, and the location of the destination marker. Each robot's orientation was varied four ways: directly pointing forward, right, left, or backward. The

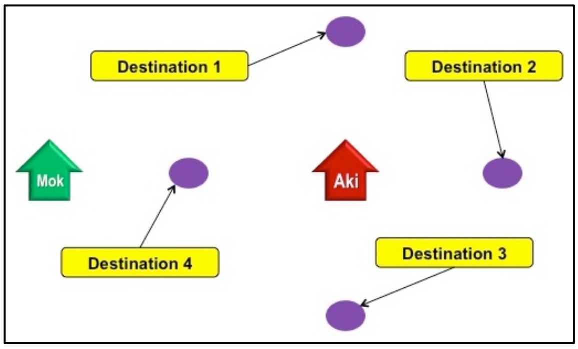

Figure 2. Specified are four potential goal destinations for Mok, the actor in all scenarios. Only one of the destinations is shown on a particular trial.

destination marker was also varied four ways: directly in front of, behind, right of, or left of Aki. These three dimensions were varied using a factorial design, yielding 64 different configurations that were presented in randomized order. Thus each participant produced 64 sets of instructions. Participants received a break at the halfway point of the session.

3.2 Schematic (2-D) Scene Experiment

Participants observed two-dimensional configurations of schematics that contained two robots (Mok and Aki) and a destination marker in this experiment. Each participant viewed a single monitor displaying a recording interface overlaid by static slides that contained the stimuli. After each participant was shown the speech recording interface and had tried it out, they proceeded through a randomly ordered slide set. In this experiment, participants viewed an overhead perspective of the scene, with the robots represented as arrows and the destination marked by purple circles (see Figures 1a and 2). The robots were represented by arrows that were meant to indicate their orientations in the scene.

3.3 Virtual (3-D) Scene and Distance Awareness Variation Experiment

In this experiment, we crafted stimuli with a three-dimensional map builder and USARSim, a virtual simulation platform designed for conducting experiments with robots (Carpin et al., 2007). The map was designed such that trials were "rooms" in a multi-room environment. Participants did not walk through the environment; they only viewed static configurations. Included in the map were instances of two Pioneer P2AT robots. All visual stimuli were presented at an eye-level view, with eyes at a height of 5'10" (see Figure 1b). The room was designed such that walls would be too far away to serve as landmarks. Visual stimuli for this experiment required full-screen access to the game engine, so the recording interface was moved to an adjoining monitor.

We included an additional condition: informing participants (or not) of the distance between the two robots. We recruited fourteen participants for this study, seven in each of two conditions. In one condition (*no-dist*), participants were not given any information related to the scale of the robots and area in the stimuli. This is equivalent to what participants experienced in the schematic scene experiment. In the second condition (*dist*), the instructions indicated that the two robots, Mok and Aki, were seven feet apart. However, no scale information (e.g., a ruler) was provided in the scene itself. This would provide the option to cast instructions in terms of absolute distances. The option to use Aki as a landmark reference point remained the same as in the first experiment. We hypothesize that participants that are not given a sense of scale will use landmarks much more often than those participants that are provided distance information.

3.4 Real-World Scene Experiment

In natural environments, it can be assumed that people generally have a good sense of scale. In this experiment, participants viewed similar stimuli to the virtual scenarios (eye-level view), but in-person (see Figure 1c). Bins were used to represent the two robots, with two eyes placed on top of each bin to indicate orientation. As in the previous experiments, participants were told to give instructions to one robot (Mok) so that it would arrive at the destination. We recorded participant instructions for 8 different configurations of the two robots (destination varied four ways, Mok's orientation varied two ways, right and left; Aki's orientation did not change). We simplified the number of orientations because we found that orientations of Mok and Aki did not influence landmark use in the previous experiments. After each instruction, participants were asked to close their eyes as the experimenter changed the orientations. Since they were not at a computer screen for this experiment, only verbal instructions were recorded, with no task times.

3.5 Participation

A total of 35 participants were recruited for this study, 10 in the schematic scene experiment, 14

Environment	Type	Spoken language route instruction (transcribed with fillers removed)
2-D	Mixed	Mok turn left / and stop at the right hand side of *Aki*.
2-D	Mixed	Turn right about sixty degrees / then go forward until you're in front of *Aki*.
3-D no-dist	Mixed	Mok turn to your left / move towards *Aki* when you are pretty close to *Aki* stop there / turn to your right / continue moving in a straight line path you will find a blue dot to your left at some point stop there / turn to your left / and reach the blue dot which is your destination.
3-D no-dist	Relative	Go forward half the distance between you and *Aki*.
3-D dist	Absolute	Rotate to your right / move forward about five feet / rotate again to your left / and move forward about seven feet.
3-D dist	Absolute	Turn to your right / move forward one foot / turn to your left / move forward ten feet / turn to your left again / move forward one foot.
Real-world	Absolute	Okay Mok I want you to go straight ahead for about five feet / then turn to your right forty five degrees / and go ahead and you're gonna hit the spot in about four feet from there.
Real-world	Mixed	Mok move to *Aki* / turn left / and move forward three feet.

Table 1. Spoken language route instructions for Mok, the moving robot, were transcribed and divided into absolute and relative steps (absolute step / relative step). Absolute steps are explicit instructions that contain metric or metric-like distances, while relative steps include Aki (the static robot) as a reference.

in the virtual scene experiment, and 11 in the real-world scene experiment. Participants ranged in age from 19 to 61 (M = 28.4 years, SD = 9.9). Of all participants, 22 were male and 14 were female. All participants were self-reported fluent English speakers.

4 Data

The first study (schematic stimuli) yielded a total of 640 route instructions (64 from each of 10 participants). All of these instructions were transcribed in-house using the CMU Communicator guidelines (Bennett and Rudnicky, 2002). In addition to the recorded instructions, we also logged participants' interactions with the speech recording interface. Since the experiment instructions ask participants to think about what they plan to say before recording their speech, we assessed their "thinking time" from this logging information.

In the second study (virtual stimuli), more participants were recruited, but they were divided into two conditions (presence/absence of an explicitly stated metric distance between the two robots in the stimuli). A total of 896 route instructions were collected in the second study (64 from each of 14 participants). Of the 14 participants recruited for this study, 12 were transcribed using Amazon's Mechanical Turk (Marge et al., 2010) with the same guidelines as the first study. In the real-world study, 8 route instructions were recorded from 11 participants and transcribed, yielding a total of 88 utterances.

5 Measurements

Several outcomes were analyzed in this study, including the time needed to formulate directions to the robot and the number of discrete steps that participants included in their instructions. We analyzed two measures, "thinking time" and word count. Thinking time represents the time between starting viewing a stimulus and pressing the "Record" button. We measured utterance length by counting the number of words spoken by participants for each instruction. Utterance-level restarts and mispronunciations were excluded from this count.

We also coded the instructions in terms of the number of discrete "steps" (see Table 1). We defined a "step" as any action where motion by Mok (the moving robot) was required to complete a sub-goal. For example, "*turn left and*

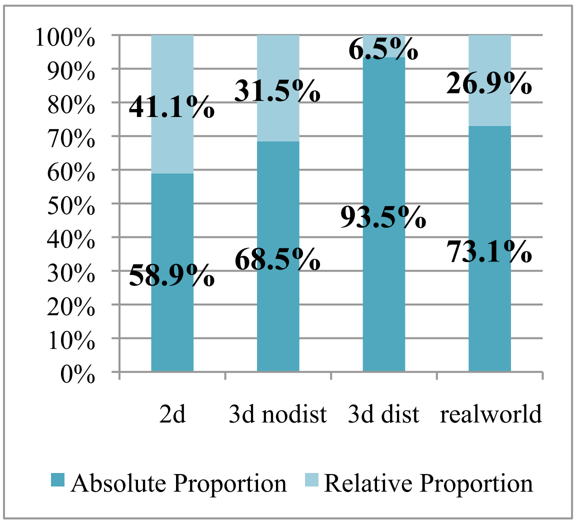

Figure 3. Mean proportion of relative steps to absolute steps across distance-naïve 2-D (schematic), distance-naïve 3-D (virtual), distance-aware 3-D (virtual), and real-world scenarios (with a 1% margin of error).

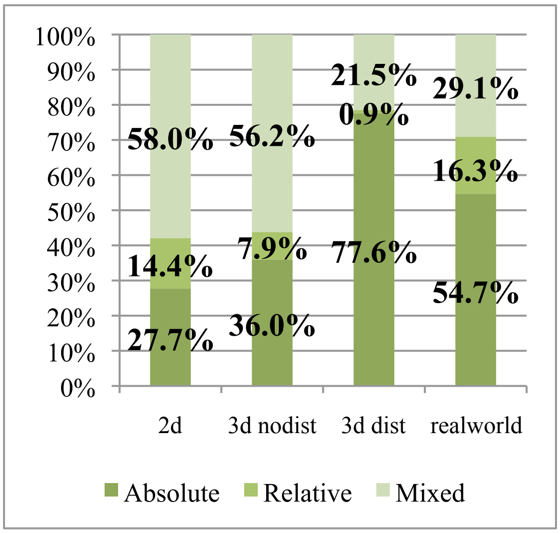

Figure 4. Proportions of instruction types across distance-naïve 2-D (schematic), distance-naïve 3-D (virtual), distance-aware 3-D (virtual), and real-world scenarios.

move forward five feet" consists of two steps: (1) a ninety degree turn to the left and (2) a movement forward of five feet to get to a new location. We divided steps into two categories, *absolute steps* and *relative steps* (similar to Levinson's (1996) *absolute* and *intrinsic* reference systems). An *absolute step* is one with explicit instructions that contain metric or metric-like distances (e.g., "*move forward two feet*", "*turn right ninety degrees*", "*move forward three steps*"). We assume that simple turns (e.g., "*turn right*") are turns of 90 degrees, and thus are absolute steps. We define a *relative step* as one that includes Aki, the static robot, in the reference (e.g., "*move forward until you reach Aki*", "*turn right until you face Aki*").

6 Results

We conducted analyses based on measures of thinking time, word count, and the number of discrete "steps" in participants' spoken language route instructions. Among the folds of the data we examined were observations from schematics without distance information (i.e., "*2-D no-dist*"), virtual scenes without giving participants distance information (i.e., "*3-D no-dist*"), virtual scenes with giving participants initial distance information (i.e., "*3-D dist*"), and real-world scenes (i.e., "*realworld*"). Since we collected an equal number of route instructions in the two virtual scene conditions (i.e., with and without being told about the distance in the environment), we directly compared properties of these instructions.

In Sections 6.2 and 6.3, absolute steps, relative steps, word count (log-10 transformed), and thinking timing (log-10 transformed) were the dependent measures in mixed-effects models of analysis of variance (for significance testing). ParticipantID was modeled as a random effect. We are interested in the population from which participants were drawn.

6.1 Adjusting Spatial Information

Landmark use was affected by participants' awareness of scale. The fewer scale cues available, the greater the number of references to landmarks. Thus, landmarks were most prevalent in instructions generated for schematic scenarios and least prevalent in the condition that explicitly specified a scale. See Figure 3 for the actual proportions. We did not inform participants of scale in the real-world condition. Interestingly, their absolute/relative mix was closer to the no-scale conditions even though they were observing an actual scene and could presumably make inferences about distances. Figure 4 shows that presentation format also affected participants' use of instructions that were entirely absolute in nature. There were fewer mixed instructions (i.e., instructions where absolute instructions were supported by landmarks) in conditions where participants had a sense of scale.

Though distances may be self-evident in real-world scenarios, they often are not in virtual environments. Participants behaved differently from real-world scenarios when we presented a non-trivial indication of scale. Participants' instructions were dominated by absolute instructions when they had a sense of scale in a virtual environment. This suggests that despite similarities in scale awareness, people formulate spatial language instructions differently when they cannot for themselves determine a sense of distance in an environment.

6.2 Sense of Distance in Virtual Stimuli

We directly compared participants' spoken language route instructions with respect to the presence (i.e., "*dist*") or absence (i.e., "*no-dist*") of distance information in the virtual environment. Though participants already had an initial preference toward using metric-based instructions, these became dominant when participants were aware of the distance in the virtual environment.

Participants that were not given a sense of distance referred to Aki as a landmark much more than when participants were given a sense of distance, confirming our initial hypothesis. We observed that the mean number of relative steps in the *no-dist* condition was nearly four times greater (1.0 relative steps per instruction) than the *dist* condition (0.2 relative steps per instruction) ($F[1, 12] = 4.6$, $p = 0.05$). As expected, participants used absolute references more in the *dist* condition, given the lack of landmark use. The mean number of absolute steps was greater in the *dist* condition (3.3 per instruction) compared to the *no-dist* condition (mean 2.4 absolute steps per instruction) ($F[1, 12] = 5.5$, $p < 0.05$).

As shown in Figure 3, the proportions of absolute to relative steps in participants' instructions show clear differences in strategy. When participants received distance information, an overwhelming majority of steps were absolute in nature (i.e., steps containing metric or metric-like distances). Aki was mentioned in steps only 6.5% of the time in the dist condition (i.e., relative steps). The proportions were more balanced in the *no-dist* condition, with 68% of steps being *absolute*. The remaining 32% of steps referred to Aki. The difference between proportions from the *no-dist* and *dist* conditions was statistically significant ($F[1,12] = 7.5$, $p < 0.05$). From these analyses we can see that distance greatly influenced participants' language instructions in virtual environments.

We further classified participants' instructions as entirely absolute, relative, or mixed in nature. When participants used landmarks, they tended to mix them with absolute steps in their instructions. Participants in the *dist* condition comprised most instructions with only absolute steps. However, even though 6.5% of steps were absolute in nature, they were distributed among one-fifth of instructions. In the *no-dist* condition, though relative steps comprised only 31.5% of total steps, they were distributed among a majority of the instructions. These results suggest that sequences of absolute steps may be sufficient on their own, but relative steps, when used, depend on the presence of some absolute terms.

6.3 Goal Location and Orientation Results

Our analysis showed that the goal location in scenarios impacted participants' instructions. For word count, participants used significantly different numbers of words based on the goal location ($F[3, 1580] = 252.2$, $p < 0.0001$). Upon further analysis, across all experiments, when the goal was closest to the Mok, the moving robot, people spoke fewer words (14 fewer words on average) compared to other locations (analysis conducted with a Tukey pairwise comparisons test). Participants also had significantly different thinking times based on the goal location ($F[3, 1502] = 6.21$, $p < 0.05$). Thinking time for the destination closest to Mok was lowest overall (on average at least 1.3s lower) and significantly different from two of the three remaining goal locations (via a Tukey pairwise comparisons test). There were no significant differences in word count and thinking time when varying Mok's orientation or Aki's orientation.

We also observed patterns in the steps people gave in their instructions. A landmark's placement, when directly interfering with a goal, increased its reference in spatial language instructions. When the goal location was blocked by Aki, we observed a high proportion of relative steps. For schematic stimuli, participants often required Mok to move past Aki in order to get to the destination. After observing the proportions of absolute steps and relative steps out of the total number of steps across destination, we found that stimuli with this destination yielded an average of 45% relative steps to 55% absolute steps. This is a greater proportion than any of the other destinations (their relative step proportions ranged from 33% to 38%).

7 Summary and Conclusions

We presented a study that examines people's close-range spoken language route instructions for robots and how the presentation format and the complexity of the route influenced the content of instructions. Across all presentation formats, people preferred providing instructions that were absolute in nature (i.e., metric-based). Despite this preference, landmarks were used on occasion. When they were, participants' use of them was influenced by the presentation format (schematic, virtual or natural). When participants had a general sense of distance in scenes, they were much more acclimated to using specific distances to give route instructions to a robot.

Our results indicate that the goal location can influence participant effort (i.e., time to formulate) and the pattern (absolute/relative) in spoken language route instructions to robots. Several of these were predictable (e.g., least effort when goal location was closest to moving robot). When participants viewed these configurations in virtual environments, there were clear differences in their instructions based on whether or not they were given a sense of scale.

We compared the natural language instructions from the real-world condition to those from virtual stimuli. Figure 3 shows that in general, real-world participants' instructions contained similar proportions of landmarks to the *3d no-dist* (virtual) condition. However, there was a greater preference to use absolute steps in the real-world than in the virtual world; participants apparently access their own sense of scale when formulating these instructions. With respect to spatial language instructions, participants tended to treat virtual environments much like real-world environments.

This study provides useful information about methodology in the study of spatial language and also suggests principles for the design of spatial language understanding capabilities for robots in human environments. Specifically, virtual world representations, under suitable conditions, elicit language similar to that found under real-world situations, although the more information people have about the metric properties of the environment the more likely they are to use them. But even in the absence of unambiguous metrics people seem to want to use such language in the instructions that they produce. These observations can be used to inform the design of spatial language understanding for robot systems as well as guide the development of requirements for a spatial reasoning component.

Acknowledgments

This work was supported by the Boeing Company and a National Science Foundation Graduate Research Fellowship. The authors would like to thank Carolyn Rosé, Satanjeev Banerjee, Aasish Pappu, and the anonymous reviewers for their helpful comments on this work. The views and conclusions expressed in this document only represent those of the authors.

References

C. Bennett and A. I. Rudnicky. 2002. *The Carnegie Mellon Communicator Corpus*, ICSLP, 2002.

G. Bugmann, E. Klein, S. Lauria, and T. Kyriacou. 2004. *Corpus-based robotics: A route instruction example*, Intelligent Autonomous System, pp. 96-103.

L. A. Carlson and P. L. Hill. 2008. *Processing the presence, placement and properties of a distractor during spatial language tasks*, Memory and Cognition, 36, pp. 240-255.

S. Carpin, M. Lewis, J. Wang, S. Balakirsky, and C. Scrapper. 2007. *USARSim: A Robot Simulator for Research and Education*, International Conference on Robotics and Automation, 2007, pp. 1400-1405.

M. P. Daniel and M. Denis. 2004. *The production of route directions: Investigating conditions that favour conciseness in spatial discourse*, Applied Cognitive Psychology, 18, pp. 57-75.

W. G. Hayward and M. J. Tarr. 1995. *Spatial language and spatial representation*, Cognition, 55 (1), pp. 39-84.

A. Klippel and S. Winter. 2005. *Structural Salience of Landmarks for Route Directions*, COSIT 2005, pp. 347-362.

T. Kollar, S. Tellex, D. Roy, and N. Roy. 2010. *Toward Understanding Natural Language Directions*, Human Robot Interaction Conference (HRI-2010), pp. 259-266.

T. Koulouri and S. Lauria. 2009. *Exploring Miscommunication and Collaborative Behaviour in Human-Robot Interaction*, SIGdial 2009, pp. 111-119.

S. C. Levinson. 1996. *Frames of reference and Molyneux's question: cross-linguistic evidence*, in P. Bloom, M. Peterson, L. Nadel, and M. Garrett (Eds.), Language and space, pp. 109-169.

K. Lovelace, M. Hegarty, and D. R. Montello. 1999. *Elements of good route directions in familiar and unfamiliar environments*, in C. Freksa and D. M. Mark (Eds.), Spatial information theory: Cognitive and computational foundations of geographic information science. Berlin: Springer.

M. MacMahon. 2007. *Following Natural Language Route Instructions*, Ph.D. Thesis, University of Texas at Austin.

M. MacMahon, B. Stankiewicz, and B. Kuipers. 2006. *Walk the Talk: Connecting Language, Knowledge, and Action in Route Instructions*, 21st

National Conf. on Artificial Intelligence (AAAI), 2006, pp. 1475-1482.

M. Marge, S. Banerjee, and A. I. Rudnicky. 2010. *Using the Amazon Mechanical Turk for Transcription of Spoken Language*, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2010. Dallas, TX.

P. E. Michon and M. Denis. 2001. *When and why are visual landmarks used in giving directions?* in D. R. Montello (Ed.), Spatial information theory: Foundations of geographic information science, pp. 292-305. Berlin: Springer.

R. Moratz, T. Tenbrink, J. Bateman, and K. Fischer. 2003. *Spatial knowledge representation for human-robot interaction*, Spatial Cognition III. Berlin: Springer-Verlag.

C. Nothegger, S. Winter, and M. Raubal. 2004. *Selection of salient features for route directions*, Spatial Cognition and Computation, 4 (2), pp. 113-136.

D. Perzanowski, D. Brock, W. Adams, M. Bugajska, A. C. Schultz, and J. G. Trafton. 2003. *Finding the FOO: A Pilot Study for a Multimodal Interface*, IEEE Systems, Man, and Cybernetics Conference, 2003. Washington, D.C.

M. Raubal and S. Winter. 2002. *Enriching wayfinding instructions with local landmarks*, in M. J. Egenhofer and D. M. Mark (Eds.), Geographic information science, pp. 243-259. Berlin: Springer.

R. Ross. 2008. *Tiered Models of Spatial Language Interpretation*, International Conference on Spatial Cognition, 2008. Freiburg, Germany.

H. Shi and T. Tenbrink. 2009. *Telling Rolland where to go: HRI dialogues on route navigation*, in K. Coventry, T. Tenbrink, and J. Bateman (Eds.), Spatial Language and Dialogue (pp. 177-190). Oxford University Press.

M. Skubic, D. Perzanowski, A. Schultz, and W. Adams. 2002. *Using Spatial Language in a Human-Robot Dialog*, IEEE International Conference on Robotics and Automation, 2002, pp. 4143-4148. Washington, D.C.

E. Stopp, K. P. Gapp, G. Herzog, T. Laengle, and T. Lueth. 1994. *Utilizing Spatial Relations for Natural Language Access to an Autonomous Mobile Robot*, 18th German Annual Conference on Artificial Intelligence, 1994, pp. 39-50. Berlin.

L. Talmy. 1983. *How language structures space*, in H. Pick, and L. Acredolo (Eds.), Spatial Orientation: Theory, Research and Application.

C. Vorwerg. 2009. *Consistency in successive spatial utterances*, in K. Coventry, T. Tenbrink, and J. Bateman (Eds.), Spatial Language and Dialogue. Oxford University Press.

E. Weissensteiner and S. Winter. 2004. *Landmarks in the communication of route instructions*, in M. Egenhofer, C. Freksa, and H. Miller (Eds.), GIScience. Berlin: Springer.