Evaluating Effects of Hand Pointing
by an Image-Based Avatar
of a Navigation System

Michiko Inoue[1], Aya Shiraiwa[2], Hiroki Yoshimura[2], Masashi Nishiyama[2,3(✉)],
and Yoshio Iwai[2]

[1] Graduate School of Sustainability Science, Tottori University,
101 Minami 4-chome, Koyama-cho, Tottori 680-8550, Japan
`s132007@eecs.tottori-u.ac.jp`
[2] Graduate School of Engineering, Tottori University, 101 Minami 4-chome,
Koyama-cho, Tottori 680-8550, Japan
`nishiyama@eecs.tottori-u.ac.jp`
[3] Cross-informatics Research Center, Tottori University, 101 Minami 4-chome,
Koyama-cho, Tottori 680-8550, Japan

Abstract. We investigate whether the behavior of pointing at a map
by an image-based avatar helps a user understand a route in an image-
based avatar navigation system. We also evaluate whether this behavior
is preferred by the user. Existing avatar-based methods inform the user of
a route by this behavior while talking. However, the existing methods do
not consider how to incorporate a map. Thus, we consider how to inform
the user of a route using an image-based avatar that indicates the route
by pointing at a map. In the experiments, after users interacted with
the system, we conducted a route depiction test to determine whether a
user was able to correctly understand the route on a map and performed
a questionnaire-based subjective assessment to determine whether the
user liked the image-based avatar system. The results of the experiments
show that the pointing behavior significantly increased the likeability of
the system but did not help the user understand the route.

Keywords: Navigation system · Hand pointing · Image-based avatar
Route depiction test · Questionnaire

1   Introduction

An interaction system with a life-size display has many potential applications. In
particular, there is a demand for a system that uses an image-based avatar [1–5]
to smoothly communicate with users. The avatar provides good usability, such
when talking about past experiences [1] or acting as a guide at a museum [2].
This paper discusses a route navigation system that uses an image-based avatar
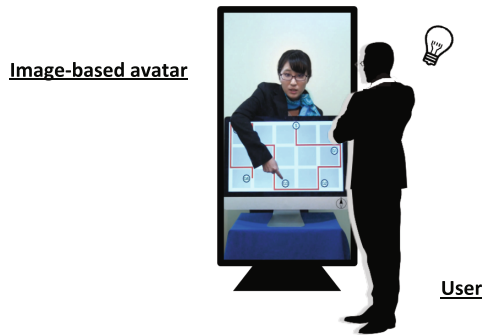for intuitive interaction, that is, as if a real guide were directing the user, as

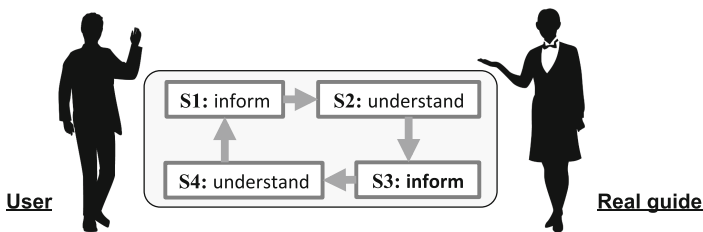**Fig. 1.** Route navigation system that uses an image-based avatar that points at a map.

Fig. 2. Cycle of route navigation as performed by people.

illustrated in Fig. 1. We assume the scenario of an information center in a public space, such as a tourist information office.

In the design of a route navigation system [6–9], the aim is for the user to understand the explanation of the route and like using the system. When a user cannot understand the route, he or she will repeatedly ask about the route and then feel uncomfortable using the system. To avoid this problem, we need to consider the interface between the image-based avatar and user.

When designing a user-friendly interface for route navigation, we aim to mimic the behaviors of a real guide. The real guide generally directs the user according the following steps.

S1: The user informs the real guide of the destination.
S2: The real guide understands the destination provided by the user.
S3: The real guide informs the user of the route.
S4: The user understands the route provided by the real guide.

The user's understanding and liking of the system are determined in the cycle of informing and understanding, as illustrated in Fig. 2. In this cycle, S3 is important in terms of smoothly satisfying the demands of the user. We thus focus on developing an interface for S3 using an image-based avatar.

In existing avatar-based methods [10–12] the avatar informs the user of the route by pointing while talking from a first-person viewpoint. The use of the

hand for pointing has merit in that it resembles natural communication among people. However, existing methods have not considered incorporating a map with a bird's-eye view. As described in [13], a map is an important part of helping a user to understand a route. A real guide frequently presents a route by indicating it on a map. We thus tackle the challenging issue of how to control an image-based avatar so that it presents the route on a map.

To this end, we investigated the hypothesis that an image-based avatar that indicates a route by pointing at a map helps the user understand it. We also investigated the hypothesis that an image-based avatar that points is preferred by the user. After exposing users to the system, we conducted a route depiction test to determine whether they correctly understood a route on a map, and then performed a questionnaire-based subjective assessment to determine whether they liked the system.

The rest of this paper is organized as follows. Section 2 describes our experimental design, Sect. 3 presents the results of the route depiction test, Sect. 4 describes the results of the questionnaire-based assessment, and our concluding remarks are given in Sect. 5.

2 Experimental Design

2.1 Overview

We assumed that a user accesses the route navigation system in an information center. We explored a scenario in which the user would like to visit some destinations in a particular order in a downtown area. We evaluated the effect when the image-based avatar describes the route on a map using its finger. Twenty-four participants (20 males, four females, average age 21.9 years) participated in the study. The details of our experimental design are described below.

2.2 Route on the Map

We generated a fictional map containing 3×5 square blocks. Figure 3 shows examples of the routes on the map. A route consists of a start point, destination points, and path segments. We randomly set the start point and the destination points during the experiment for each participant. We used six destination points so that the user would not easily remember them. In general, a human can remember 4 ± 1 items in short-term memory time [14]. We believe that using six destination points is a valid way to keep participants from easily getting full marks in the route depiction test. Furthermore, we set the paths that connect the destination points so that they did not cross each other. Note that we fixed the number of the corners in the paths to 12 to keep the experimental conditions the same.
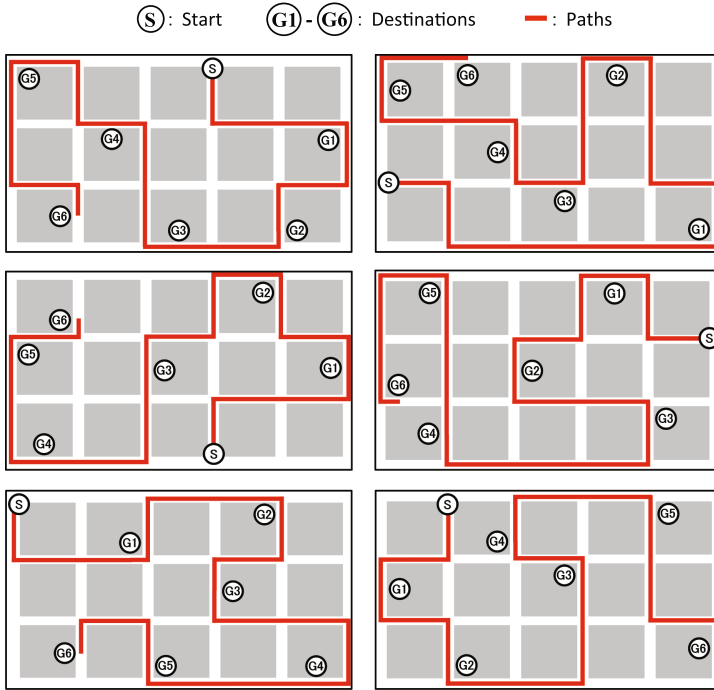
Fig. 3. Examples of the routes, which consist of a start point, destination points, and path segments on the map.

2.3 Interface with the Image-Based Avatar

We tested two interfaces as follows:

I1: An image-based avatar *with* map pointing,
I2: An image-based avatar *without* map pointing.

We generated the video sequences of the interfaces, as illustrated in Fig. 4. Each participant viewed the video sequence of interface I1 or I2 in random order. The length of each video sequence was 90 s. The sentences and speed of the avatar's speech were the same for both I1 and I2.

Figure 5 shows the setup of the interface using the image-based avatar. Each participant stood 1.5 m from the display and viewed the video sequences. We used an 80-inch display with a resolution off $1,920 \times 1,080$ pixels (Sharp PN-A601) to show the life-sized avatar. We placed the voice speaker (Towa electronic TW-S7B) behind the display.
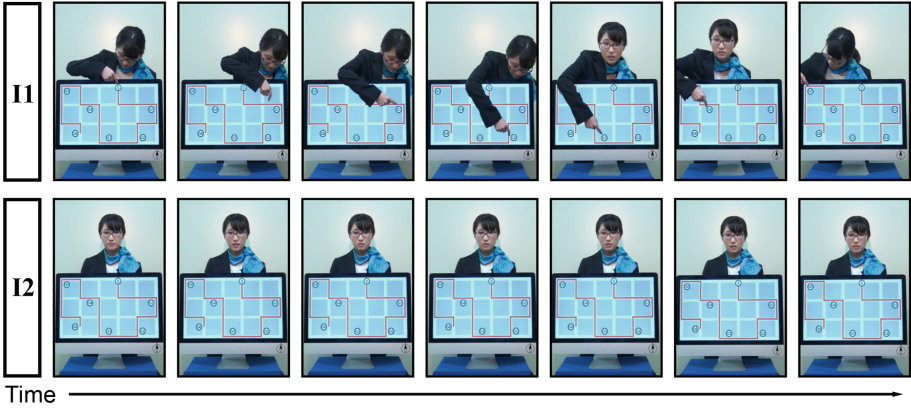
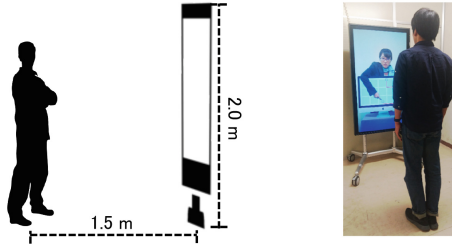**Fig. 4.** Video sequence of the interface using the image-based avatar with and without map pointing.

Fig. 5. Setup of the interface using the image-based avatar.

2.4   Procedure

To evaluate the hypothesizes for the interfaces, we executed the following procedure:

P1: We displayed the video sequence of the interface for the participant.
P2: We gave the route depiction test to the participant.
P3: We gave the questionnaire to the participant.

We also gave an easy numerical calculation task to the participant between P2 and P3. We prepared six routes on the map and randomly selected a route when assessing the interface. The order of interface I1 or I2 was randomly chosen.

3   Route Depiction Test

3.1   Overview

We prepared a blank map for the participant, as illustrated in Fig. 6(a), before starting the route depiction test. We slightly shifted the viewpoint of this blank map with respect to the map displayed in the video sequences of Fig. 4. We asked

the participants to depict the start point, destination points, and path segments at the same scale as displayed in the video sequence. Figures 6(b) and (c) show the results of the route depiction test with respect to the ground-truth of the route illustrated in Fig. 6(d).
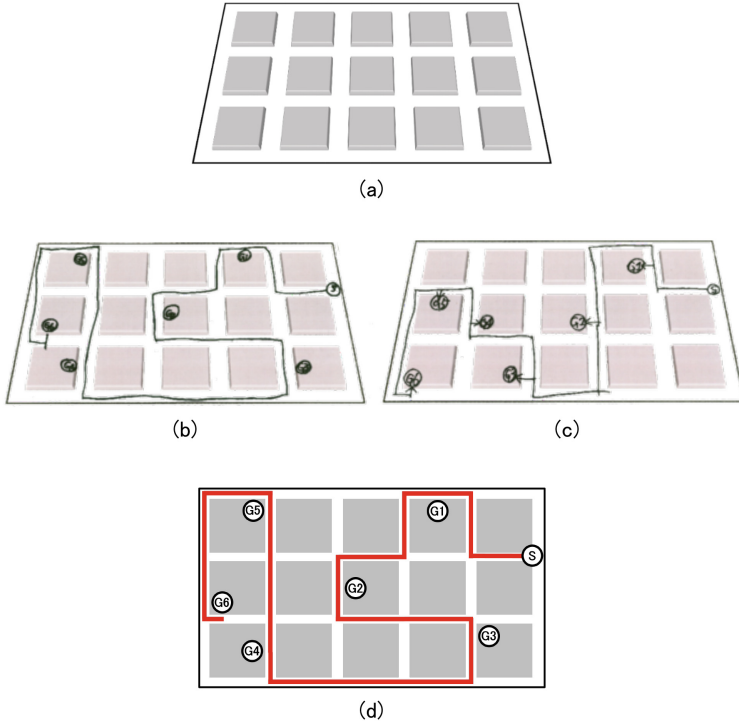

Fig. 6. Examples of the route depiction test. (a) Blank map. (b) and (c) Results of the route depiction test. (d) Ground-truth.

We evaluated the interfaces using the following three metrics. The first one was the correctness of the start point and destination points. The second one was the correctness of the path segments. The third one was the time taken for the user to complete the depiction. The details of the metrics are described below.

3.2 Metrics in the Route Depiction Test

We first explain the correctness of the start point and destination points. We evaluated whether a point depicted by the participant in the test was at the same location as the point displayed in the video sequence. When checking this correctness, we divided the blocks of the map into $3 \times 3 = 9$ regions. Figure 7(a) shows an example of the regions in a block. The depicted point was considered correct when it was more than half-way within the same region as the

displayed point. Figure 7(b) shows an example of a correct case, and Fig. 7(c) shows an example of an incorrect case, where the depicted point is shifted by one region. The maximum number of incorrect points was seven (one starting point and six destination points).
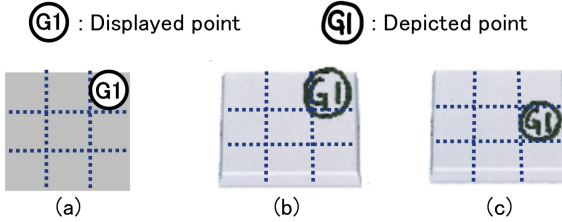

Fig. 7. Example of correct and incorrect destinations in the route depiction test. (a) Regions in the block. (b) Correct case. (c) Incorrect case.

We next explain the correctness of the path segments. We assigned correctness to a path segment when the depicted segment and the displayed segment were the same. Figure 8(a) shows an example of a displayed path segment, Fig. 8(b) shows an example of a correct case, and Fig. 8(c) shows an example of an incorrect case, where the path between the destination points was incorrect even though the locations of the points were correct. The maximum value of path incorrectness was six.
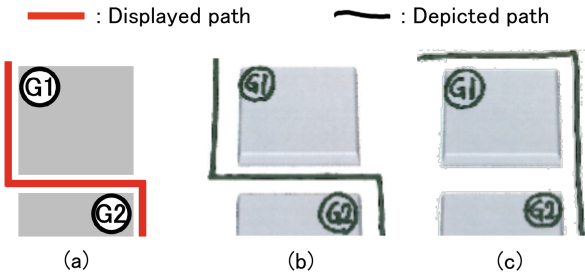

Fig. 8. Example of correct and incorrect path segments in the route depiction test. (a) Displayed path. (b) Example of a correct case. (c) Example of an incorrect case.

We finally explain the time taken by the user to create the depiction. We used a stopwatch to record the times when the participant started the test and when he or she finished.

3.3   Results of the Route Depiction Test

Figure 9(a) shows the number of incorrect start and destination points for each interface. Figure 9(b) shows the number of the participants obtaining an incorrect
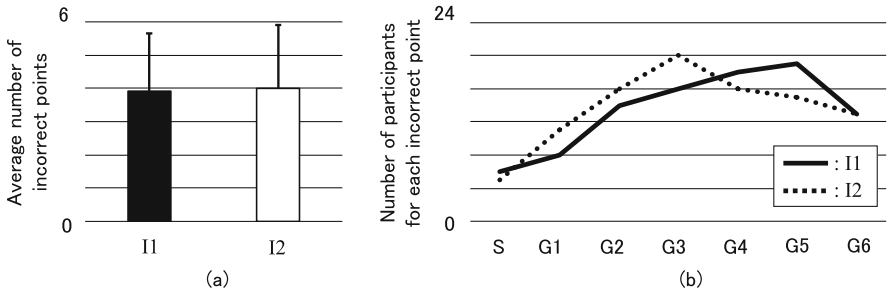
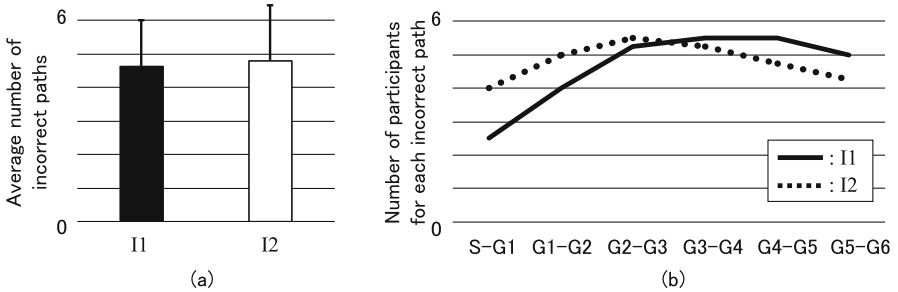**Fig. 9.** Results for destination points in the route depiction test.

Fig. 10. Results for path segments in the route depiction test.

answer for each point. We used the Wilcoxon signed-rank test to determine that there is no significant difference ($p < .05$) between interfaces I1 and I2.

Figure 10(a) shows the number of incorrect paths for each interface. Figure 10(b) shows the number of the participants obtaining an incorrect path for each path segment. We used the Wilcoxon signed-rank test and determined that there is no significant difference ($p < .05$) between interfaces I1 and I2.

Figure 11 shows the time taken to depict the route. We used the Wilcoxon signed-rank test and found that there is no significant difference ($p < .05$) between interfaces I1 and I2.
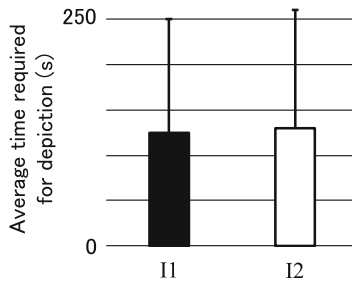

Fig. 11. Results of the average time taken to depict the map.

Hence, we cannot claim that the pointing behavior of the image-based avatar helps the user understand the route.

4   Questionnaire-Based Subjective Assessment

4.1   Items of the Questionnaire

After viewing the video sequences of interface I1 or I2, we asked the participant the following questions:

Q1: Did you like interacting with the avatar?
Q2: Did the avatar provide a navigation service resembling that of a real guide?
Q3: Was it easy to understand where to go on the map?

Each participant provided a rated score using six response levels (1: disagreeable to 6: agreeable) for each question. We also asked the inverse questions of Q1, Q2, and Q3. The purpose of Q1 was to evaluate the hypothesis that the image-based avatar that points to the map was more liked by the user. The purpose of Q2 was to check whether the behavior of the avatar in the interface was close to that of a real guide. The purpose of Q3 was to check whether the user felt that the avatar that pointed at the map had correctly presented the locations on the route.
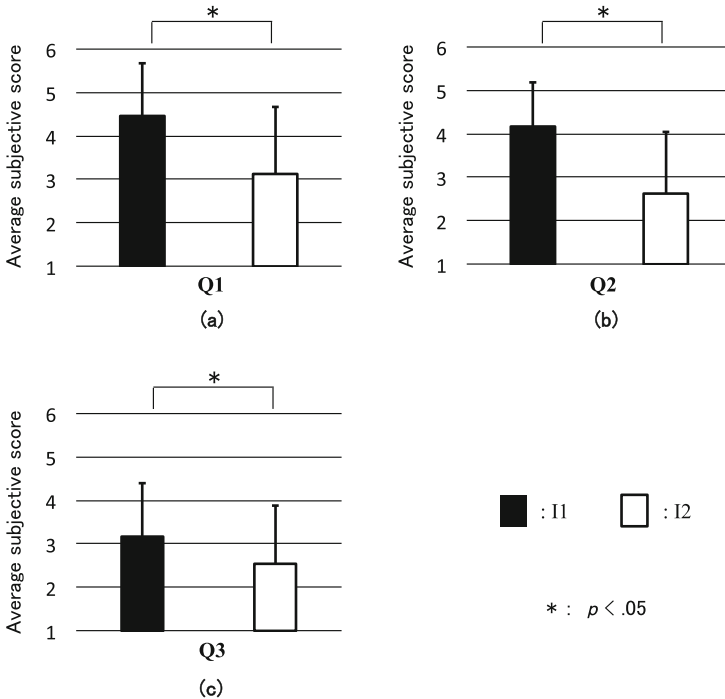

Fig. 12. Subjective scores of the questionnaire-based subjective assessment.

4.2   Results of the Subjective Assessment

Figures 12(a) to (c) show the subjective scores of the questionnaire obtained using the Wilcoxon signed-rank test. In Fig. 12(a), in terms of Q1, there was a significant difference between I1 and I2. We can hence claim that the image-based avatar that points at the map is more liked by the user. In Fig. 12(b) in terms of Q2, there was a significant difference between I1 and I2. We can hence claim that the behavior of the avatar in interface I1 is closer to that of a real guide. In Fig. 12(c), for Q3, there was also a significant difference between I1 and I2. Therefore, we can also claim that the avatar with pointing behavior makes the user feel that the avatar has correctly presented the locations on the route.

5   Conclusions

We investigated two hypotheses regarding an interface with an image-based avatar that points at a map. We evaluated the interface using a route depiction test and a questionnaire-based subjective assessment. We can claim that an avatar that points at a map significantly increases the likeability of the system, but we cannot claim that this avatar better helps the user to understand the route.

In future work, we will expand our assessment of the interactive system and intend to develop a method to add an explanation of landmarks, as used in [15], for the our avatar system.

References

1. Artstein, R., Traum, D., Alexander, O., Leuski, A., Jones, A., Georgila, K., Debevec, P., Swartout, W., Maio, H., Smith, S.: Time-offset interaction with a holocaust survivor. In: Proceedings of the 19th International Conference on Intelligent User Interfaces, pp. 163–168 (2014)
2. Robinson, S., Traum, D., Ittycheriah, I., Henderer, J.: What would you ask a conversational agent? Observations of human-agent dialogues in a museum setting. In: Proceedings of the Sixth International Conference on Language Resources and Evaluation, pp. 28–30 (2008)
3. Nishiyama, M., Miyauchi, T., Yoshimura, H., Iwai, Y.: Synthesizing realistic image-based avatars by body sway analysis. In: Proceedings of the Fourth International Conference on Human Agent Interaction, pp. 155–162 (2016)
4. Jones, A., Unger, J., Nagano, K., Busch, J., Yu, X., Peng, H.I., Alexander, O., Bolas, M., Debevec, P.: An automultiscopic projector array for interactive digital humans. In: ACM SIGGRAPH 2015 Emerging Technologies
5. Miyauchi, T., Ono, A., Yoshimura, H., Nishiyama, M., Iwai, Y.: Embedding the awareness state and response state in an image-based Avatar to start natural user interaction. IEICE Trans. Inf. Syst. E100.D **12**, 3045–3049 (2017)
6. Darken, R.P., Peterson, B.: Spatial orientation wayfinding and representation. Design, Implementation, and Applications. Handbook of Virtual Environments. CRC Press, Boca Raton (2002)

7. Devlin, A.S., Bernstein, J.: Interactive wayfinding: use of cues by men and women. J. Environ. Psychol. **15**(1), 23–38 (1995)
8. Blades, M., Spencer, C.: How do people use maps to navigate through the world? Cartographica: Int. J. Geogr. Inf. Geovis. **24**(3), 64–75 (1987)
9. Makimura, Y., Yoshimura, H., Nishiyama, M., Iwai, Y.: Decreasing physical burden using the following effect and a superimposed navigation system. In: Virtual, Augmented and Mixed Reality, pp. 533–543 (2017)
10. Kopp, S., Tepper, P.A., Ferriman, K., Striegnitz, K., Cassell, J.: Trading spaces: how humans and humanoids use speech and gesture to give directions. In: Conversational Informatics: An Engineering Approach, pp. 1–26 (2007)
11. Bergmann, K., Kopp, S.: Gnetic - using Bayesian decision networks for iconic gesture generation. In: Proceedings of the 9th International Conference on Intelligent Virtual Agents, pp. 76–89 (2009)
12. Bergmann, K., Kopp, S.: Knowledge representation for generating locating gestures in route directions. In: Proceedings of Workshop on Spatial Language and Dialogue, pp. 1–13 (2005)
13. Taylor, H.A., Tversky, B.: The description of routes: a cognitive approach to the production of spatial discourse. Cah. Psychol. Cogn. **35**, 371–391 (1996)
14. Cowan, N.: The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav. Brain Sci. **24**(1), 87–114 (2001)
15. Denis, M.: The description of routes: a cognitive approach to the production of spatial discourse. Cahiers Psychologie Cogn. **16**(8), 409–458 (1997)