The effect of a reference object’s orientation on the apprehension of spatial terms

Takatsugu Kojima • Takashi Kusumi

Keywords  Spatial cognition • Spatial language • Orientation of a reference object

Background

The interaction between spatial cognition and language is important to both cognitive science and artificial intelligence. For example, if an interactive humanlike agent could understand spatial terms in natural language in a human manner, simulating human behavior and cognitive processes should be possible. Research in artificial intelligence and linguistics has provided theoretical and computational models of the apprehension of spatial language (Gorniak and Roy 2004; Herskovits 1986; Kreizer 1997; Yamanashi 2000). Cognitive psychology has empirically examined such theoretical and computational works and provided empirical data on the relationship between spatial cognition and language.

Imai et al. (1999) examined the assignment of four Japanese spatial terms, mae, ushiro, hidari, and migi (similar to front, behind, left, and right, respectively, in English), in three-dimensional (3-D) computer graphics (CG) space. They found that the four spatial terms did not categorize the given space equally, and that the categorical patterns of space by these spatial terms were influenced by the orientation of the reference; object. When a reference object had no inherent front, mae/ushiro and hidari/migi had symmetrical boundaries in a pair, and the spatial areas of mae/ushiro were larger than those of hidari/migi. When a reference object had an inherent front, the areas of the four terms varied according to the orientation of the reference object. When a reference object was facing toward or away from participants, hidari/migi boundaries were asymmetrical but mae/ushiro boundaries were asymmetrical. When a reference object was facing sideways, mae/ushiro and hidari/migi had asymmetrical boundaries as pairs. Imai et al. (1999) concluded that the orientation of a reference object influenced the apprehension of the spatial terms dynamically in 3-D space. However, they compared only the boundaries of the four terms and did not examine how the effect influenced the spatial categorical pattern of each spatial term. Their method had some limitations and was not capable of fully examining the effect of a reference object’s orientation. Furthermore, no detailed data were provided for applications to humanlike agents. Thus, we examined the effect of each of the same four terms in 3-D space, with the goal of providing detailed experimental data and general implications for spatial cognition and language understanding.

Method

Sixty native Japanese graduate and undergraduate students at Kyoto University participated in this study. We divided the participants into three groups of 20 people each, with each group assigned one of the following orientation conditions: rotations of 0°, 90°, and 180°. The experiment was run on a computer with a 17-in. monitor. In the experiment, two windows appeared on the screen. The display window appeared at the top of the screen and was used to present stimuli, whereas the evaluation window appeared below the display window and was used to present statements and receive responses. Four Japanese projective spatial terms were used: mae, ushiro, hidari, and migi (in front of, behind, to the left of, and to the right of in English, respectively). These are basic Japanese spatial terms for 3-D space. The sentences used in this experiment were “Akai tama-wa midori-no hako-no xx-ni aru” (“The red ball is xx the green box”). One of the four spatial terms was assigned to xx.
The stimuli were constructed using Open GL and were presented in a 3-D CG space in the display window. All the stimuli contained the ground, a red ball (the located object), and a green box (the reference object). We attached a yellow part to the green box and instructed the participants that the yellow indicated the front of the box. The ground was divided into a 7 x 7 matrix that was invisible to the subjects. The ground surface was covered with a texture map that looked like a random dot land surface because we did not want the participants to use any ground-surface cues.

The positions of the reference object (the green box) and the viewpoint were fixed at respective coordinates in the 3-D CG space. We fixed the frame of reference under each orientation condition. The participants were instructed to note the specific frame of reference before they participated in the experiment. All of the objects were presented at the centers of the ground cells.

Under each orientation condition, each location was used only once for each spatial term. Therefore, each condition consisted of 108 trials (four spatial terms x 27 locations). Each participant viewed the computer screen from a distance of approximately 60 cm in a dark room, and was instructed about the frame of reference according to the condition (the 0°, 90°, or 180° rotation condition). In each trial, a cross-shaped fixation point appeared at the center of the display window, and a statement and a button to indicate a stimulus were presented in the evaluation window. The participants read the sentence, clicked on the button, and viewed a scene in the display window after a 1,000-ms delay. After the scene was presented, subjects judged whether the statement matched the scene by pressing a computer mouse button (a left button for “yes” and a right button for “no”) as quickly as possible. In the experiment, there was no correct answer for whether the statement matched the scene. However, we instructed the participants to choose a correct answer in the yes/no task so that we could collect reaction time (RT) data.

Then, nine rating buttons, which represented a scale ranging from 1 (least applicable) to 9 (most applicable), appeared in the evaluation window. The participants judged how well the statement matched the scene, chose one of the nine buttons by moving the mouse-pointer, and pressed the button. In each experiment, the participants were given ten practice trials, after which they participated in 108 experimental trials. The order of presentation was generated randomly for each participant. The participants were allowed to rest between trials.

**Results**

We analyzed the mean acceptability ratings. For each spatial term, a two-way 3 (orientations: 0°, 90°, and 180° rotations) x 27 (27 locations) mixed analysis of variance (ANOVA) of the acceptability rating in each location was performed at a significance level of .05 with orientation as a between-subject variable and location as a within-subject variable.

For mae, significant main effects were found for both the orientation, F(2, 57) = 4.64, P = .0135, and the location, F(26, 1482) = 204.58, P < .0001. A significant interaction was also determined, F(52, 1482) = 2.13, P < .0001. As a post hoc test for the main effect of orientation, we conducted a Ryan’s test for a significance level of .05. Significant differences were revealed in the combinations of 0° and 180° rotations (P = .0064) and 90° and 180° rotations (P = .0200). Because the orientation had significant simple main effects, we performed a Ryan’s test for a significance level of .05.

For ushiro, significant main effects were observed for both the direction, F(2, 57) = 6.78, P = .0023, and the location, F(26, 1482) = 325.85, P < .0001. A significant interaction was also found, F(52, 1482) = 3.11, P < .001. As a post hoc test for the main effect of the orientation, we conducted a Ryan’s test for a significance level of .05. Significant differences were revealed in the combinations of 0° and 180° rotations (P = .0019) and 90° and 180° rotations (P = .0045). As the orientation had significant simple main effects, we conducted a Ryan’s test for a significance level of .05. For hidari, the tests revealed no significant main effect of orientation but did find a significant effect of location, F(26, 1482) = 199.97, P < .0001. A significant interaction was determined, F(52, 1482) = 1.55, P < .0077. Because the orientations showed significant simple main effects, we conducted a Ryan’s test for a significance level of .05.

For migi, significant main effects were found for both the orientation, F(2, 57) = 3.57, P = .0346, and the location, F(26, 1482) = 198.21, P < .0001. A significant interaction was also found, F(52, 1482) = 1.36, P = .0472. The orientations had significant simple main effects. As a post hoc test for the main effect of orientation, we conducted a Ryan’s test for a significance level of .05. Significant differences were revealed in the combinations of 0° and 90° rotations (P = .0030) and 0° and 180° rotations (P < .0001). Because orientation showed significant simple main effects, we conducted a Ryan’s test for a significance level of .05.

We then analyzed the mean RT in the same way. For each spatial term, however, no significant main effects or interaction effects were observed. The results of the mean acceptability ratings indicated that the orientation of the reference object influenced the spatial categorical patterns of the four spatial terms in the 3-D CG space. For example, when a reference object was facing away from participants (under the 180° rotation condition), the spatial areas indicated by mae or ushiro were larger than under the 0° rotation condition.
In contrast, when a reference object was facing sideways (under the 90° rotation condition), the spatial areas indicated by hidari or migi were narrower than under the 0° rotation condition. The mean RTs did not have any significant main effects or interaction effect for each spatial term. This result indicated that reference frame selection and adjustment (Logan and Sadler 1996; Carlson-Radvansky and Logan 1997) were experimentally controlled in this experiment.

Conclusions

These results suggest that our apprehension of spatial language is visually and dynamically influenced in 3-D space and demonstrate the inadequacy of computing the apprehension of spatial language using only the spatial positions and geometric properties of objects. The rotation of a reference object may influence the apprehension of projective spatial terms in 3-D space. Visual cues of 3-D information might be processed during spatial template construction. Our results also provide more detailed data that can be applied to the development or improvement of humanlike agents capable of understanding Japanese spatial terms in 3-D space.

References