

Title	参照枠と参照点構造の理論により空間表現文の指示対象の曖昧性を解消する対話システムの研究
Author(s)	宮城, 友香
Citation	
Issue Date	2021-06
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/18033">http://hdl.handle.net/10119/18033</a>
Rights	
Description	橋本敬, 先端科学技術研究科, 修士(知識科学)

## Abstract

Spatial language sentences in dialog systems are often ambiguous. This ambiguity results from the location of target objects which can be interpreted in several ways based on the frame of reference. In this study, using the theory of cognitive linguistics, we aimed at formulating a method of disambiguation in a dialogue system for Japanese spatial language sentences. Specifically, we disambiguated object referred by utterances by asking subjects follow-up questions based on a flowchart that specifies two parameters: whose point of view they are viewing the space from, and how they perceive the object. As a result, we were able to disambiguate the instructional object within a certain range.

The ambiguity of spatial language sentences comes in many forms. For example, suppose we ask a robot for help with moving packages; "Put the box in front of the TV." There are several destination locations. Depending on interpretation, putting the box in front of the direction that the TV is facing and putting the box between the robot and the TV are both appropriate choices. Resolving such ambiguities is a major challenge in current dialogue systems.

In existing research, the dialogue system disambiguates the target of the instruction without further clarification, resulting in inconsistencies between the user's and system's interpretations. We investigate and propose a method to resolve the ambiguity of the target of spatial language expressions in dialog using the theory of cognitive linguistics. The theory of cognitive linguistics takes the position that the result of the cognitive subject's interpretation of the environment through the body appears as a linguistic expression. As mentioned earlier, since the ambiguity of the target of instruction is caused by the interpretation of the cognitive subject, it is effective to use the theory of cognitive linguistics.

First, we investigated the reasons for ambiguity of the object of instruction in spatial language, based on the theories of reference point structure and reference frame in cognitive linguistics. We found that the object of instruction in spatial sentences changes depending on two parameters: from whose point of view the space is viewed, and how the object is perceived. In other words, ambiguity was caused by the existence of multiple viewpoints and perceptions.

Next, we investigate whether there is ambiguity in spatial language sentences in actual daily situations. Specifically, we showed native Japanese speaker participants a picture of a table with multiple objects on it. They were asked to make a spatial language sentence targeting one of the objects. 26.4% of the collected sentences had ambiguity of

the indicated object. These sentences cannot be disambiguated unless the point of view and how the object was perceived are clarified.

Based on the results, we devised flowchart for disambiguating the indicated objects in spatial language. It incorporates the theories of reference point structure and frame of reference in cognitive linguistics, and conditionally branches and narrows down on which viewpoint is taken and how the object is perceived.

Finally, we verified the effectiveness of the flowchart proposed. We built a dialog system that allows users to communicate with a robot in a three-dimensional virtual space through text. The native Japanese speaker participants were asked to input sentences to the robot, where they request that the robot pick up an object specified. Upon receiving the text input from the participants, the experimenter interpreted the participant's spatial language requests based on the flowchart. If necessary, the participants were asked follow-up questions via text. As a result, 18.1% of the sentences contained ambiguity, all of them were successfully resolved, and the intended object identified. In total, we revealed that 62.5% of expressions took the frame of reference of the robot. It was hypothesized that, due to participants making requests to the robot, the viewpoint of the person receiving the request was given priority. It was also noted that the perception of space changes depending on the shape of the target object.

In this study, we propose a method for resolving the ambiguity of the target of instruction in spatial language that occurs in dialogue systems. By asking the participants follow-up questions based on a flowchart which identify two parameters: from whose point of view the space is viewed, and how the object is perceived, we resolved the ambiguity and successfully identified the instruction's intended target in all cases. However, the experimental environment of this study is limited in scope. Application of the system in real world environments is limited due to the increase in complexity. Therefore, it is necessary to incorporate techniques such as image processing and focal chaining in which reference point structures are chained multiple times.