Addressee-dependent Blending in Object Localisation Tasks*

Kerstin Fischer

University of Bremen – kerstinf@uni-bremen.de

1 Blending in Spatial Reference

In this study, cognitive operations will be investigated for their use in spatial object localisation tasks involving different addressees. The cognitive operations considered are conceptual metaphor (1), metonymy, and blending (2). Blending concerns the integration of two independent representations into a new structure and is particularly relevant in our scenario.

The first pervasive finding in our data is that as soon as three similar objects are close to each other, speakers refer to them as a group (3; 4). For instance, reference to an object of this group is achieved by referring to it as the left, middle, or right object, irrespective of whether they are right, middle or left from the participants. Thus, the *right object* may well be left of speaker or hearer. We can understand this as a blending operation of the concept of a group with the close spatial proximity of the objects in question, in comparison with other objects. The group is then taken as the relatum for spatial reference.

Furthermore, if the placement of the objects 'resembles' the shape of an object, speakers refer to the groups in terms of these objects, for instance, *the square, the triangle, the top row, the bottom line*. That is, objects were conceptualised as points marking the corners of imaginary objects. Yet, they are not *treated* as 'making up' these corners, but as parts of real objects, as can be seen in the following utterance: "the brick is *located* at the top corner of the triangle."

Blending of different conceptual spaces occurs at several levels. Especially the up/down dimension (even though our settings are two-dimensional, that is, objects are placed on the ground in front of the participants) can be found pervasively, for instance, "top row of group of three" or "bottom of square". According to (5), the relation between *up* as referring to a path on the vertical dimension and *up* as a goal-directed path is metonymical: the reaching of its full extension in the vertical dimension is taken to stand for the reaching of a goal in general, as in *drive up the street.* In a next step, the metonymic extension of *up/down* as a path is extended to a whole region, such that the blended geometrical figures have a top and a bottom. Thus, we find the metonymic treatment of *up* and *down* for *far* and *close* regions, as well as a blending of object locations with geometrical figures. The latter is in line with reference to, for instance, Tangram figures in conversation (6).

* This research has been supported by the German Research Foundation in the framework of the SFB/TR8 *Spatial Cognition* at the University of Bremen.

2 Addressee-dependent Cognitive Operations

The cognitive operations for spatial reference just described are quite complex. The question I would like to address now is whether speakers can be shown to be aware of this complexity and whether they adapt their instructions to communication partners who may be suspected to have problems with such conceptualisations. The example investigated here is human-to-robot communication (HRI).

2.1 Data

The set of human-robot dialogues analysed here were elicited in a scenario in which the users' task was to instruct a robot to measure the distance between two objects out of a set of seven. These objects differed only in their spatial position. The users typed instructions into a notebook, having an overview of the objects to be referred to and the robot (a joint attention scenario). The relevant objects were pointed at by the instructor of the experiments. There were 21 participants from all kinds of professions and with different experience with artificial systems. The robot's output was either "error" (or a natural language variant of it) or a distance in centimeters. Since by reformulating their utterances the users display their hypotheses about the functioning of the system (see 7), error messages were frequent. The user utterances are typed and thus transcription was not necessary; typos were not corrected.

The human-to-human dialogues used were elicited with the same object configuration. The task was to describe one of the objects so that the addressee could identify it by pointing at it. To match the fact that in the HRI dialogues users typed their utterances, participants in this experiment were instructed to write their instructions on a piece of paper, which was then handed to the addressee. Participants were placed in the same positions as in the human-robot scenario. The participants were not allowed to talk or to communicate non-verbally.

2.2 Blending in Human-Robot Interaction

Many behaviours of speakers in human-computer and human-robot interaction show that users believe their artificial communication partners to be restricted in particular ways, especially regarding natural language processing, perception, and human abilities such as identifying the intended meaning of a sentence in spite of spelling errors. Thus, we may expect that users, in case they suspect non-literal language use and blending operations to be complex, will also adapt to their artificial interlocutor by using simpler or different types of descriptions (7).

What can be found is that speakers indeed employ these cognitively complex spatial descriptions, but not to same amount as in the human-to-human setting. Furthermore, as soon as there are problems, speakers change to another way of referring to the objects in question. As the repair in the example below shows, users are uncertain whether the notion of a group, here 'collection', can be processed by an artificial system, and thus in case of problems change to some other reference system:

(1) usr15-6:messe distanz zwischen den inneren beiden Tassen der rechten Tassenansammlung [**measure distance between the two inner cups of the right collection of cups**]
sys:ERROR 652-a: input is invalid.
usr15-7:messe distanz zwischen der tasse in der Mitte und der Tasse schräg rechts dahinter [**measure distance between the cup in the middle and the cup diagonally right behind**]

The same holds for blended geometrical figures. In the next example, the blend is marked as an imagined representation, not as a real object, and even though the system signals understanding by providing a distance, the speaker employs another kind of description which does not make use of cognitively complex concepts, such as an imaginary rectangle:

(2) usr17-9:miß den Abstand zwischen den beiden Tassen, die die längere Seite eines Rechtecks darstellen würden [**measure the distance between the two cups that would represent the longer side of a rectangle**]
sys:31,1 cm
usr17-10:miß den Abstand zwischen Roboter und der Tasse, die am zweitnächsten steht [**measure the distance between robot and the cup that is second next**]

Moreover, even the metonymic *up/down* dimension which can be seen as highly conventionalised is hardly used at all. In contrast to the pervasive usage in the human-to-human situation, only five out of twenty-one users in the HRI distance measurement task employed *up* and *down* at all. These users are furthermore likely to change to another way of instructing, as in the example below:

(3) usr13-30:miss die entfernung zwischen dem objekt links oben und dem objekt rechts oben [**measure the distance between the object up left and the object up right**]
sys:Die Instruktion konnte nicht erkannt werden. Bitte formulieren Sie neu. [**The instruction could not be processed. Please reformulate.**]
usr13-31:miss die entfernung zwischen dem objekt links hinten und dem objekt rechts hinten [**measure the distance between the object back left and the object back right**]

3 Conclusion

We can conclude that the speakers' ideas about the cognitive complexity of spatial descriptions influences the strategies they take in the communication with their artificial communication partners. Speakers are aware of potential problems in the use of cognitively complex ways of talking about space by means of blended spatial expressions, such as *the top row of the square*, and they adapt the complexity of their instructions to the (supposed) capabilities of their communication partners.

Bibliography

[1] Lakoff, G., Johnson, M.: Metaphors We Live by. The University of Chicago Press (1980)

[2] Fauconnier, G., Turner, M.: The Way we Think. Basic Books (2002)

[3] Moratz, R., Fischer, K., Tenbrink, T.: Cognitive modelling of spatial reference for human-robot interaction. International Journal on Artificial Intelligence Tools **10** (2001) 589–611

[4] Tenbrink, T., Moratz, R.: Group-based spatial reference in linguistic human-robot interaction. In: Proceedings of EuroCogSci 2003: The European Cognitive Science Conference, September 10-13, Osnabrück, Germany. (2003) 325–330

[5] Lindner, S.J.: A Lexico-Semantic Analysis of English Verb-Particle Constructions with *out* and *up*. PhD thesis, University of California (1983)

[6] Clark, H.H., Wilkes-Gibbs, D.: Referring as a collaborative process. Cognition **22** (1986) 1–39

[7] Fischer, K.: Linguistic methods for investigating concepts in use. In Stolz, T., Kolbe, K., eds.: Methodologie in der Linguistik. Frankfurt a.M.: Peter Lang (2003)