

Conveying spatial information in linguistic human-robot interaction

Thora Tenbrink

FB10: Faculty of Linguistics and Literary Sciences

University of Bremen, Germany

tenbrink@sfbtr8.uni-bremen.de

Abstract

An overview is presented of the range of variability employed by speakers with regard to the specificity of information in spatial communication, focussing on the significance of the context. This variability is described in terms of two underlying dimensions, one of which reflects a dichotomy of underdeterminacy and redundancy, while the other concerns vagueness and precision. Relevant results of several HRI experiments are used for exemplification.

1 Introduction

How does the given context influence the richness and unambiguousness of information or detail in spatial communication? While it is well established that redundancy is pervasive in linguistic communication, an impressive amount of literature deals with the ways in which utterances can be underdetermined or vague, establishing that most language is neither precise nor unequivocal. Linguistic spatial communication is specifically problematic. In any discourse, many aspects of context contribute to understanding the processes enhancing or hampering the success of the communicative goals. No matter how much theoretical ambiguity exists for an utterance, in most cases speakers manage to convey enough information for a human listener to infer what is meant. Modelling these processes in order to enable automatic language understanding systems to achieve similar communicative success is a major challenge for research in NLP.

Experiments in human-robot interaction offer an optimal research area in this regard, since the dis-

course context (including linguistic elements such as the robot's output and the discourse history as well as situational factors such as spatial configurations) can be controlled and modified to an exceptionally high degree. Leaving the test persons ignorant with respect to the research aims, the density and specificity of information conveyance can be tackled on the basis of natural language data where the circumstances of production are specifiable in great detail.

One useful distinction (Pinkal 1985) adopted in the present analysis is that between *vague* utterances that leave room for a *continuous* spectrum of precisification options, and *ambiguous* ones that offer *discrete* options for precisification. While Pinkal addresses these phenomena principally on the lexical level, they can occur on all levels of linguistic communication, such as grammatical, pragmatic, or conceptual.

2 Method

Natural language data were collected in several studies of human-robot interaction of two different kinds. In each study, human users were asked to instruct an unfamiliar robot to perform a task with regard to several objects placed on the floor together with the robot. The tasks to be performed were either to move to one of the objects, or to measure the distance between two of them. The robot variably occupied different positions. The configurations differed with regard to complexity, and while in the 'move' tasks the robot understood specific kinds of utterances (which the users were not told about), the 'measure' tasks were faked. Here, the computer into which the users typed was not connected to the robot, but produced answers in a predetermined fashion. Thus, different kinds of 'dialogues' emerged. Together, the studies con-

stitute a data pool of natural language instances of spatial reference in human-robot interaction scenarios in which (in contrast to most other work in spatial human-robot interaction) the users were not informed about the robot's functionalities.

3 Results and Discussion

Utterances in which some piece of information necessary for identification of the goal object is missing and needs to be inferred are classified as *underdetermined*, while in cases of *redundancy* some information is provided in more than one way. These phenomena occurred in the following areas in our data:

- **Perspective.** Since there are three kinds of possible perspectives, some specification is necessary but mostly not provided by users.
- **Reference systems.** Some spatial expressions are capable of reflecting different kinds of reference systems: 'rechts' can be used in relative and intrinsic reference systems, which in some configurations may be a source for ambiguity.
- **Directions and angles.** Some linguistic expressions (e.g. 'diagonal') presuppose a vector, needing a starting point as well as an end point for interpretation. Utterances containing less information are therefore underdetermined. Utterances indicating an angle but leaving out information about the underlying reference frame are underdetermined, as in "minus 10 Grad" (minus 10 degrees).
- **Configuration and figures.** Some expressions presuppose knowledge about the spatial setting for interpretation. Implicit reference to groups (as in the adjectival use of projective terms) belongs in this category as well as all expressions containing superlatives ('farthest', 'nearest'). Other expressions reflect the conceptualisation of a specific figure, such as a triangle or a square. Utterances containing no explicit definition of such a group or figure are underdetermined in this regard.
- **Linguistic underdeterminacy.** Viewed in isolation, all utterances containing phoric (i.e., anaphoric, cataphoric or exophoric) elements are inherently underdetermined, since they depend on the – situational or textual – context and require the recipient to infer the intended referents. In spatial communication, it is natural to refer (deictically) to the perceived spatial

surroundings, and exophoric elements are therefore pervasive in the data. For successful communication it is essential that the interlocutors' perception is compatible or can be matched via the linguistic representations.

Often, users failed to provide information with regard to one of these categories, while giving elaborate and redundant information on a different aspect. This variability may reflect the user's current focus of attention in the given interaction.

On another dimension, instructions vary between *vagueness* and *precision*. Many linguistic expressions specify spatial relations in a 'qualitative' fashion, i.e., do not require the speaker to provide explicit information about quantitative measures. Overwhelmingly, in the data the qualitative information given by spatial expressions was not specified further. Hedges and modifications as well as combinations of several reference systems were used to indicate the speakers' awareness of the vagueness of their utterances, as in "das vordere Objekt etwas links von der Mitte" (the front object a little to the left of the middle). Thus, users attempt to render the instruction more precise by narrowing the range of applicability. This occurred most often in more complex scenarios.

The classic difference between 'continuous' and 'discrete' kinds of specification possible in cases of indefiniteness is reflected in spatial language and can be applied to conceptual as well as linguistic phenomena. Underdetermined utterances contain expressions presupposing elements that need to be inferred by the context, while vague utterances need not be presuppositional, but contain inherently 'fuzzy' expressions that represent spatial phenomena qualitatively rather than quantitatively.

The phenomena identified in our data can be interpreted via the notion of **contrast**: Users' choices on the scales between redundancy and underdeterminacy as well as vagueness and precision reflect their aim at referring to the goal object(s) in a way that is sufficiently distinct to any competing objects, rendering the interpretation highly dependent on a reliable mapping between the user's and the system's conceptualisations of the current scenario.

Reference

- Pinkal, M. 1985. Logik und Lexikon: Die Semantik des Unbestimmten. Berlin: de Gruyter.