

## **Spatial strategies in the description of complex configurations**

Thora Tenbrink<sup>1</sup>, Kenny R. Coventry<sup>2,3</sup>, & Elena Andonova<sup>1</sup>

<sup>1</sup>SFB/TR8 Transregional Spatial Cognition Research Centre

Bremen University, Germany

<sup>2</sup>Cognition and Communication Research Centre,

Northumbria University,

Newcastle, UK

<sup>3</sup>Hanse Institute for Advanced Studies,

Delmenhorst, Germany

### **Address for Correspondence:**

Thora Tenbrink

FB 10, Universität Bremen

Postfach 330440,

28334 Bremen, Germany

E-mail: tenbrink@uni-bremen.de

Tel.: +49 421-218-64212

Fax: +49 421-218-98-64212

**Running head:** Spatial strategies for complex configurations

## **Spatial strategies in the description of complex configurations**

### **Abstract**

How people describe complex arrangements of objects in a small-scale setting has not been investigated sufficiently to predict shifts or stability of discourse strategies. In a study involving 100 native German participants, we investigated speakers' choices of perspective as well as location and orientation information when describing arrangements of doll's furniture. Results point to a dominant preference for relative reference frames using an external, observer-based perspective, independent of functional aspects in the arrangement but influenced by object orientation. Objects were described either in relation to another object in the arrangement or in relation to the arrangement as a whole, with approximately equal preference. The speaker's current or imagined position was seldom used for reference in location descriptions, although it served regularly as a basis for describing object orientation. Generally, description features in our study differed from earlier studies addressing individual object-to-object or large-scale descriptions, demonstrating the impact of scale, configuration, and discourse task on speakers' choices.

(156 words)

### **Author Contributions**

The order of authorship is arbitrary; all authors contributed equally to the work reported.

Imagine you are expecting visitors, and a friend is helping you lay the table with many different items. How do you explain to her where the objects should be placed on the table, relative to one another? Where objects are can be described from various viewpoints and using many possibilities for reference. While all objects will certainly end up *on the table* or *in front of your friend* who is laying the table, some of them should be *next to* or *behind each other*, or *at the far end of the table* from your friend's working position. Not all of these descriptions are equally useful, but they are all available for use. Moreover, object locations can be described using a constant perspective from a stable position (e.g., from a particular seat at the table), or by adapting perspectives locally depending on particular configurations (e.g., from the perspective of salient objects on the table). In this paper, we examine spatial descriptions in a table-top setting in which a complex arrangement of objects (pieces of doll's house furniture) is placed directly in front of the speaker. Our goal was to identify typical as well as possible (but atypical) strategies of spatial description for 3D objects in peripersonal (near) space, and their relationship to individual object descriptions in particular local configurations. Prior to identifying the choices participants make when describing spatial layouts, we first briefly consider previous research on spatial description, highlighting how differences in methodology may have led to differences in the types of spatial description found.

### **Spatial Description in Context**

Although spatial language research has been a hotbed of activity over the last two decades, surprisingly few studies have addressed naturalistic description of spatial arrays involving multiple objects. There is a considerable body of research that has examined how people describe individual objects when presented with only a second

object in the visual scene to be described. For such scenes it has been established, for example, that object features such as salient parts (Miller & Johnson-Laird, 1976) as well as functional relationships between objects (Carlson-Radvansky & Radvansky, 1996; Coventry, Prat-Sala, & Richards, 2001) influence how their relative locations are described. Also, the presence and actions of a dialogue partner (Tversky & Hard, 2009) and the nature of the communication task as well as the prepositional inventory of a language (Grabowski & Weiss, 1996; Mainwaring, Tversky, Ohgishi, & Schiano, 2003) affect speakers' choices. However, as soon as a sequence of spatial descriptions is required (successively describing pairs of objects), more general strategies come into play that do not emerge with individual object location descriptions (Herrmann & Grabowski, 1994). For example, speakers tend to be consistent in their successive descriptions (Ehrich, 1985; Vorwerg, 2009) at least as long as no further factors influence the choice of conceptual perspective (Tversky, Lee, & Mainwaring, 1999).

Less is known about how people describe spatial arrays involving multiple objects. Several studies have used complex arrays, but have asked participants to describe the locations of individual objects in those arrays, rather than the arrays as a whole (Plumert, Carswell, de Vet, & Ihrig, 1995; Plumert, Spalding, & Nichols-Whitehead, 2001; Schober, 1995). These studies have produced some interesting findings. For example, speakers choose a perspective that is helpful for their interaction partner (Schober, 1995). Plumert and colleagues found that individual object descriptions are embedded in systematic conceptual hierarchies, and that functionally meaningful links between individual as well as encompassing objects regularly influence object localization strategies. However, these studies do not allow analyses of how participants naturally describe complete, complex arrays – including how they sequence their descriptions, to what extent they include orientation

information (the chair points to the left), and whether they choose to describe individual objects (chair, another chair, dining table, etc.) or clusters of objects (e.g., dining area; table and chairs). These aspects may play a major role in natural discourse situations such as room descriptions for prospective tenants or when suggesting interior design solutions to an absent friend, and the like.

The studies that have asked participants to freely describe complex arrangements have employed memory methodologies, where participants freely recall arrangements known to them (e.g., their own apartments or newly learned arrangements, such as maps). Of these studies several involve recall of objects in *vista*, *environment*, or *geographical space*<sup>1</sup> (see Montello, 1993 for a detailed classification of spatial scales). For example, Linde and Labov (1975), Ullmer-Ehrich (1979), and Shanon (1984) all asked participants to describe the rooms in which they lived, requiring them to mentally move around their rooms given that not all objects were visible from a single point (*vista/environmental space*). In other studies, such as those of Taylor and Tversky (1996), participants were asked to describe how to get from one place to another in an open (city) environment (*environmental space*).

While such studies are revealing regarding the perspectives participants adopt for spatial description, there is emerging evidence to suggest that spatial language is

---

<sup>1</sup> *Figural space* is “projectively smaller than the body” (Montello, 1993, p. 315), and does not require movement of the body to perceive objects in that space. *Vista* space is projectively as large or larger than the body but can be visually apprehended from a single point in space without movement of the perceiver. This includes the space of small rooms, town squares, horizons, etc. *Environmental space* surrounds the body, and can’t be perceived without locomotion on the part of the perceiver. It relates to spaces containing large buildings and cities. Although not directly perceivable in a short time, Montello argues that this level of space is nevertheless perceptible with enough exposure. Finally, *geographical space* is at a larger scale again, not directly perceptible through direct experience/navigation, and learnable only through symbolic representations, such as map formats that reduce geographical space to figural space.

not invariant across scales (Lautenschuetz, Davies, Raubal, Schwering, & Pederson, 2007). For example, English and German use cardinal directions, such as *north* and *east*, in environmental and geographical space, but not in figural space.<sup>2</sup> So there is reason to believe that spatial description for large scale spaces, where it is necessary to move around an array in order for objects to be visible, may be very different from spatial description for objects visible from a single viewpoint (Grabowski, 1999).

Moreover, there is now robust evidence of a neurological distinction between near space and far space depending on whether objects are close enough to be touched or beyond arm's reach (see for example Berti & Rizzolatti, 2002; Làdavas, 2002). This distinction maps onto the distinction between figural and vista space in Montello's classification, and has already been shown to affect spatial language choice. Coventry, Valdés, Castillo and Guijarro-Fuentes (2008) found that this distinction corresponded to switches in spatial demonstrative choice to describe object locations. Such findings indicate that even a change from figural to vista space may be crucial for complex spatial description.

Only a few studies have examined free spatial description in table top (figural) space. Carroll and von Stutterheim (1993) had people describe a composite object and model living room and village arrangements. They examined whether and how individual object descriptions relied on a global reference frame encompassing the complete configuration. Differences identified between English and German speakers were traced back to language-specific patterns of information organization; for example, German and English locative expressions reflect divisions in space in grammatically different ways with respect to transitivity structures, leading to distinct patterns of nominal phrase use. For instance, the German adverb *vorne* differs from

---

<sup>2</sup> In American English, compass directions may be used in figural space in a transferred sense, where *west* corresponds to *left*, see Carroll & von Stutterheim (1993).

the pragmatically equivalent English phrase *at the front* in that English can add a relatum such as *of the house*, whereas German grammatically requires a separate *spatial* constituent such as *im Haus (in the house)*. According to Carroll and von Stutterheim (1993) such differences affect the discourse strategies employed by speakers in the description of spatial configurations in fundamental ways.

In a series of studies conducted by Ehrich (1985) and Ehrich and Koster (1983), participants were asked to describe a configuration of objects in figural space arranged in groups, either as functionally and spatially close clusters, or as unrelated linear arrangements. Results indicate that speakers' description strategies are affected by object arrangement. However, the distance between objects was not controlled. Therefore it remains unclear whether the perspective shifts were due to the *spatial* arrangement or rather to the *functional* relationship between the objects.

All of the studies thus far using complex object arrangements in figural space have required speakers to recall the spatial arrangements from memory. However, there is good reason to believe that this task may be very different from describing spatial arrangements while they are still visible and accessible in near space. It is well established that memory for spatial relations violates formal principles of Euclidean geometry, such as axioms of symmetry (i.e., that the distance from A to B is the same as the distance from B to A). For example, distances estimated between a good landmark and a poor landmark are greater when the poor landmark is the referent than when the good landmark is the referent (e.g., Newcombe, Huttenlocher, Sandberg, Lie, & Johnson, 1999; Saddalla, Burroughs, & Staplin, 1980; see McNamara & Diwadkar, 1997 for a review). Moreover, memory for distance is also affected by knowledge of the relations between objects. Hirtle and Mascolo (1986) found that objects that were functionally related tended to be remembered as being closer

together than they actually were in the scene presented at encoding (in their case, simple drawings with object labels). Memory studies also risk the production of errors of commission – recalling objects that were not in fact present in the original array. These errors were originally noticed by Bartlett (1932) who found that background knowledge contaminated memory for simple narratives.

So both the spatial scale and whether the array to be described is present or absent during spatial description are important methodological issues when one considers spatial description. For these reasons we asked participants to describe a complex array of objects that remained in full view during description in figural space. Given that strong effects of functional relations have been extensively documented for scenes involving two objects (see Coventry & Garrod, 2004 for a review), we also decided to vary the functional relations presented in the array to be described. Before outlining the study, however, a synthesis of the previous literature provides the basis for predictions we can make regarding different aspects of spatial description.

## **Choice of relatum**

Spatial location descriptions refer to the position of a particular object in focus relative to another object or area. For our present purposes the object in focus is called the *locatum*, while the object or area in relation to which the locatum is described is called the *relatum* (rather than *reference object* which has been used in the literature to refer to both). In some spatial descriptions, the relatum may be linguistically implicit, although it is always conceptually present (e.g., when someone says *to the left*, meaning *to your left*).

**Location descriptions.** A range of systematic patterns have been identified concerning preferences with respect to the choice of relatum in spatial descriptions.

According to Talmy (2000) a relatum is typically larger, more perceivable, and more stationary than the locatum (see de Vega, Rodrigo, Ato, Dehn, and Barquero, 2002 for empirical support). Ehrich (1985) points out that descriptions of room arrangements may either continuously take the object described previously as a relatum, moving from already mentioned objects to new ones (see also Herrmann & Grabowski, 1994), or they may use one constant relatum such as a salient object within an object cluster, with the choice between these strategies dependent on the particular arrangement of objects. Carroll and von Stutterheim (1993) and Shanon (1984) found that speakers regularly use an environment-based relatum (e.g., a room and its parts) for reference. Furthermore, objects may be related to the speaker's position – either the actual position or an imagined position inside the scene during a mental tour (Ullmer-Ehrich, 1979). Which of these types of relatum (objects just mentioned, speaker position, or the environment) is dominant in complex configuration descriptions is an open question.

**Orientation descriptions.** Object location descriptions are as such unspecific concerning how the described object is *oriented* with respect to the relatum. However, orientation can be conveyed by verb use (as in *standing* vs. *lying*), or it may be specified directly, for example using similar spatial terms (*right*, *left*, etc.) as in location descriptions (Zwarts, 2003). Such descriptions of orientation may lack a relatum; they may simply have a direction based on a perspective (usually that of the observer), as in *it points to the right*, or they may use other types of linguistic expressions to represent the object's orientation, as in *this table is aligned with the room as a whole*. Orientation descriptions that do refer to a relatum may describe one of the locatum's sides as pointing towards another object or entity, as in *the chair's back points towards the wall* or *the chair is facing towards me*. How speakers choose

a relatum (if any) for descriptions of object orientation has, to our knowledge, not been investigated previously; these choices may or may not correspond to choices of relata for object *location*. Generally, the extent to which speakers make reference to object orientations at all when describing complex arrangements is currently not known. Since earlier accounts in the literature typically focus on location rather than orientation, no clear predictions can be derived, although location appears to be more prominent than orientation in spatial arrangement descriptions.

## Choice of perspective

Descriptions of the relation between locatum and relatum that use terms like *left*, *right*, *in front*, and *back/behind* are based on an underlying *perspective* as these terms can be interpreted in more than one way (see Tenbrink, in press for an overview of options). Adopting terminology suggested by Kriz and Hegarty (2005), speakers may either use an *embedded* perspective imagined inside the environment, or they may adopt an *external* perspective such as their own current position when looking at the scene. These two general options encompass a number of more fine-grained possibilities as follows.

**Embedded perspective types.** First, objects with well-defined sides of their own, such as chairs with clear fronts and backs, can be used as relata independent of an observer's point of view (using an intrinsic reference frame; Levinson, 2003). This option yields an *object-based* perspective, embedded within the scene for each single description. According to Ullmer-Ehrich (1979) speakers tend to use this perspective consistently as a strategy when object clusters are arranged around an object with intrinsic sides. But such object clusters may not always be present; furthermore, relying continuously on objects' intrinsic sides requires jumping from one object

cluster to another suitable one, which may involve high cognitive costs (Tversky et al., 1999).

Second, Linde and Labov (1975) identified a *route-based* strategy for descriptions of apartment layouts. Using this strategy, speakers travel mentally through the scene and describe the objects' locations successively relative to their own imagined position. The imagined view direction changes continuously, aligned with the imagined route. If, for instance, there is a door to the right when entering the scene which the imagined traveler then heads towards, the door's position will become *in front* rather than *to the right* of the traveler. This basic strategy has often been assumed to be the default choice whenever several spatial descriptions need to be produced in sequence. Although it might seem natural to describe the various locations encountered along the (imagined) trajectory (Linde & Labov, 1975; Taylor & Tversky, 1996) there is no evidence in the literature as yet that order of object description is indeed directly dependent on perspective choice.

Third, Ehrich (1985) suggests that speakers sometimes employ a *mental tour* with additional mental rotations toward object clusters. Then objects are described in relation to each other from the observer's imagined point of view inside the scene. For example, if there are two objects located on the right wall from an observer's point of view, the observer may mentally turn towards these objects and describe them as being to the left and right of each other. However, it has not been established whether such a strategy would be integrated in complex arrangement descriptions with regular, equidistant object positions currently visible to the speaker.

**External perspective types.** First, speakers may adopt one fixed perspective from a particular viewpoint, for instance from the door (and this fixed perspective may be stated explicitly when describing rooms from memory; Ullmer-Ehrich, 1979),

or matching the speaker's earlier view on the scene (*gaze tour*, Ehrich & Koster, 1983). With such an external perspective, any two objects can be related spatially to one another, regardless of their intrinsic features, as in *the ball is to the right of the table from my point of view*; and the locatum may also be placed inside the relatum, as in *the table is in the back of the room*. With this kind of perspective, the *vertical axis* may also be used to refer to the observer's frontal view axis, particularly in German (Carroll & von Stutterheim, 1993), as in *there's a chair in the left top corner*. The frontal axis allows for two contradicting interpretations, often referred to as *mirror* (or *facing*) vs. *tandem* (or *aligned*) interpretations. In mirror interpretations, an object that is located *in front of* another object is closer to the observer than the relatum is. In tandem interpretations, the opposite is the case. For English and German speakers, the former variant is more common (Hill, 1982).

Second, in large-scale settings, speakers sometimes employ a further external option, termed *survey perspective* by Taylor and Tversky (1996). Here objects and locations are described by relying on the cardinal directions (*north, south, east, west*).

Armed with these possible uses of spatial language identified in past literature, we tested how functional and non-functional spatial relationships are naturally described by speakers in complex (real) object arrangements that are visible in peripersonal space throughout. Given the broad variety of options for spatial description choices, we set out to investigate speakers' preferred strategies using a methodology designed to test across a range of contexts and arrays of complex arrangements.

## Method

**Participants.** 100 German native speakers participated in this study for course credit or payment (59 women, 41 men; mean age 21 years, SD=5.98, age range 17-57). Most of them were university or high school students. Participants were randomly allocated to one of six description settings – three involving functional object configuration (altogether 47 participants) and three involving non-functional object configurations (53 participants). The data were collected in German (in the language examples that follow, we provide English translations that correspond to the German original as closely as possible).

**Stimuli.** We used two different spatial configurations, which enabled us to gain insight into the impact of particular features of the arrangement on speakers' systematic spontaneous choices. The two arrangements contained 15 furniture items each (using an identical set of objects), taken from a doll's house. In the 'functional' arrangement (see [Figure 1](#)) critical items were clustered into groups of objects as they would be typically found in a house. In the 'non-functional' arrangement (see [Figure 2](#)) the items were not clustered as one would expect for a house. However, for both arrangements the items were placed equidistantly. The arrangements did not contain a door, and they were always positioned in the same orientation. Each participant saw only one arrangement, which was placed in full view in front of them; they were however not allowed to manipulate the objects.

--- Insert Figures 1 and 2 about here ---

**Procedure.** The participants were asked to describe the arrangement of the furniture items so that another person could arrange the items correctly later. The instructions

were also varied to allow generalizability; in one version, participants were told that the arrangement belonged to a living-room, in another they were told the arrangement belonged to a second-hand furniture store, and in another (neutral) version no information was given about the room (see Andonova, Tenbrink, & Coventry, in press). The descriptions were recorded and later transcribed. The task lasted approximately 10 minutes.

**Coding of responses.** The transcribed monologues were segmented pragmatically into units, as is common in spoken language analysis (e.g., Sacks, Schegloff, & Jefferson, 1974; Denis, Pazzaglia, Cornoldi, & Bertolo, 1999). Each unit corresponds to one sentence (or a “possible sentence” in the sense of Selting, 2000), and typically provides information concerning one object or a cluster, if participants so chose. Each individual unit was coded as follows.

- Content of unit: orientation or location of an object; i.e., whether the unit gives information about where an object is located vs. the direction where it is pointing, yielding four mutually exclusive categories: *only location information*, *only orientation information*, *both*, and *neither*.
- Choice of *perspective*. The diversity of potential interpretations of a spatial description in the present scenarios precludes a definite assessment of perspective, particularly for those object location descriptions that use another object (rather than the room or whole arrangement) as a relatum. Since the current view on the scene (corresponding to the gaze-based strategy) is an obvious default perspective choice for speakers (supported, for instance, by Ehrich, 1985), we focused on compatibility with this option. Units that did not rely on a perspective (such as *the chair is in the middle of the room*) were treated as neutral. Units that were (potentially) compatible with other perspectives in addition to the gaze perspective were treated as compatible with gaze perspective. This yields three mutually exclusive categories: *perspective-neutral*, *compatible with gaze perspective*, and *incompatible with gaze perspective*. Units that were incompatible with the gaze-based strategy,

i.e., survey-based, or based on an embedded kind of perspective (object-, route-, or mental tour-based as described above) were further examined qualitatively, as described below.

- Choice of *relatum*. For this analysis, we further looked into the content identified in the first annotation step. To identify the relatum of location and orientation descriptions we used two categories: units that contained an object location description (i.e., units with *only* location information plus units with *both* types of information), and units that contained an object orientation description (i.e., units with *only* orientation information plus units with *both* types of information). For each of these two types of unit, we identified whether it conveyed a clearly identifiable relatum for the location or orientation description, respectively (e.g., in *the chair is to the left of the table*, *table* serves as relatum in a location description, as does *wall* in *the chair points to the wall* in an orientation description). As stated above, unlike location descriptions, orientation descriptions do not presuppose a relatum. The following mutually exclusive choices of a relatum were further coded for location as well as orientation descriptions that conveyed an identifiable relatum:
  1. *Object-based relata*: The locatum is described relative to another object as relatum (*behind the table*). These descriptions fell into two subcategories: they either relied on an object just mentioned (either explicitly or by relying on a pronominal expression such as *behind it* "dahinter"), or they referred to a different object which was not mentioned in one of the three previous units (cf. the concept of *bridges* in Shanon, 1984) and thus not in the current focus of attention. We refer to these two subcategories as *given* vs. *new*, respectively.
  2. *Environment-based relata*: The locatum is described relative to the environment or scene as a whole, referred to regularly as the *room* and sometimes by other encompassing terms (*picture*). Since the conceptualization of an encompassing relatum represents an internal spatial relationship (an object is being described as being in a position *inside* the room), the terms *vorne* (*in front*) and *oben* (*at the top*) were interpreted as belonging in this category even without explicit mention of a relatum, as these terms can only be used for internal relationships in

German (Klabunde, 1999). Furthermore, *parts* of the environment or scene (*corner, middle of room, side, wall, upper half* etc.) were also included in the category of environment-based relata; these represent external relationships.

3. *Speaker-based relata*: The locatum is described relative to the speaker (*in front of me*).
4. *Other relata*: This category contains descriptions with more than one spatial relation (*behind it in front of the chair at the wall there is a table*), complex relata (*between*), rows and figures serving as relata (*in the third row, in the triangle*) and a few other idiosyncrasies.

All codings were discussed iteratively between several coders until agreement was reached. In addition to (and as part of the process of) iterative coding, we also checked for inter-coder agreement by independent double coding for a subset of 226 (11.3%) of all units (taken from all conditions). Levels of agreement were very good or good: Cohen's kappa for coding of content (orientation and location information) was 0.83, for choice of perspective 0.81, for choice of relatum in location descriptions 0.75, and for choice of relatum in orientation descriptions 0.83.

## Results

Altogether, we collected 2073 units. Of these, 851 were produced with respect to the functional array, and 1222 with respect to the nonfunctional array. A comparison of the mean number of utterances revealed that this is a reliable difference; participants produced a greater number of utterances for the non-functional array ( $M=23.06$ ) than for the functional array ( $M=18.12$ ),  $t(98)=2.11$ ,  $p=0.04$ .

**Description content.** Speakers regularly described both the location and the orientation of objects, but considerably more information was given concerning the

former. Out of the whole corpus of 2.073 units, 61.5% (produced by all 100 speakers) contained only location and 7.8% (produced by 46 speakers) only orientation descriptions, 17.0% (produced by 56 speakers) contained both location and orientation descriptions, and 13.8% (produced by 64 speakers) contained neither type of information. Table 1 shows the mean use of location only, orientation only, both location and orientation, and neither location nor orientation descriptions by spatial array type. A 2 (content of unit: only location information, only orientation information, both, and neither) x 2 (array type: functional versus non-functional) mixed ANOVA on the mean percentages produced across participants<sup>3</sup> revealed a main effect of content of unit,  $F(3, 294) = 128.75, p < 0.0001$ , and an interaction between content of unit and array,  $F(3, 2940) = 3.94, p < 0.01$ . The pattern for content of unit was the same for the functional and non-functional array, except for a greater percentage use of both location and orientation content in the non-functional compared to the functional array ( $p < 0.001$ ).

--- Insert Table 1 about here ---

**Perspective.** Speakers described the array for the most part by relying on their own current view on the scene (*gaze perspective*), and in part by using descriptions that were neutral with respect to perspective. Of the 2.073 units in the corpus, 35.7% (produced by 97 speakers) were perspective neutral, and 59% (produced by 99 speakers) were compatible with the external gaze perspective. Only 110 (5.3% of the total amount of units in the corpus) were incompatible with this perspective. 41 of these were produced by 23 different speakers in the functional condition, and 69 were

<sup>3</sup> To deal with skewness, dependent measures for this and later ANOVAs were logarithmically transformed by a 10-base logarithm with one added,  $\log_{10}(x+1)$  to avoid the undefined  $\log_{10}(0)$ .

produced by 32 different speakers in the non-functional condition. No difference in the frequency of use of the gaze perspective incompatible descriptions was found comparing speakers describing the functional versus the non-functional array ( $t(98) = 1.35$ , n.s.; 6.0% for the functional array and 6.4% for the non-functional array).

Speakers tended to be consistent, though there were individual differences. One speaker consistently used an external *survey perspective* (partly mixed with gaze perspective), and one speaker consistently – and rather explicitly – used a consistent embedded *mental tour* based perspective, which included mental rotations towards the objects along the imagined path. Notably, the theoretical option of a globally adopted *route perspective*, which uses embedded viewpoints along an imagined route while relating the objects to the imagined traveller's position and travelling perspective, was not in evidence in the data.

The following typical example (1) illustrates the prominent gaze-based strategy (the description refers to the functional arrangement shown in Figure 1).

- (1) Okay in the front in the middle there is a table a yellow table, left and right of it two chairs, further left all the way in the corner a small low cupboard. All the way to the right in the corner there is a sofa with a cushion on it. In the top left there is a cabinet; between the cabinet and the small cupboard on the left at the wall there is a high shelf. In the right corner there are two armchairs placed at some distance from each other with two cushions. Further to the left of that in the middle top there is a smaller armchair. All the way uh further to the left of that, not quite in the corner yet there is a further cupboard. All the way to the right uhm a bit apart from the edge there is a lower yellow table, and directly in the middle there are two stools, and further to the left a chair.

In this example, the speaker exclusively uses her current view on the scene, starting in the middle of the row closest to her, then jumping from the left corner to the one on the right, back again to the left side, and so on. Some of the objects are described with respect to each other, others relate to subregions of the configuration (corners, middle,

edge). The external gaze perspective is used consistently, sometimes with the frontal axis and sometimes with the vertical.

The following example refers to the nonfunctional arrangement ([Figure 2](#)):

- (2) In the top left corner there is a sofa with a cushion. In front of that or below it there is a shelf. Below that in the left corner there is a, how do you call this thing again, an armchair with a cushion on it, then on the top right beside the sofa there's a chair. Below that, roughly at the level of the shelf, there's another chair, and then at the bottom, at the level of the armchair, a small table, a half-round, no, quarter of a circle. Then on the top right again beside the chair there's a kind of armchair, that's a strange couch element or so; below that there is again a chair and on the bottom a table. Again, on the top right, there's a shelf, a kind of wall shelf with three levels; below that another quarter-circle table, and all the way down a small shelf with two levels. In the top right corner there is an armchair with a cushion, in the middle there's a table and on the bottom a sideboard, and that is all the way to the right in the corner.

This speaker also uses the external gaze perspective consistently, moving from one object to the next in a fairly regular fashion. Both of these examples are provided in the original German version together with their annotations in Table 2.

--- Insert Table 2 about here ---

However, not all descriptions were as consistent as these two examples. Most (94 cases; i.e., 85.5%) of the 110 descriptions that were not compatible with the gaze perspective occurred in location descriptions for a single specific set of objects, namely, the set situated on the lefthand wall in either one of the two arrays (plus, in the nonfunctional version, the chair on the remote wall close to them), as in the following example (cf. [Figure 2](#)):

- (3) All the way to the left at the wall [*gaze*] there is in the middle a big shelf, uhm to the right of which [*embedded*] is a double sofa. That is the larger sofa uhm it also stands at the wall. To the left of the big shelf [*embedded*] there's an armchair.

[*German original*: ganz links an der Wand steht in der Mitte ein großes Regal, ähm rechts davon ein Zweiersofa; also das größere Sofa das hat ähm also das steht an der Wand auch dranne. Links von dem großen Regal steht ein Sessel]

This speaker refers to a shelf positioned on the left wall, using the *gaze* perspective.

The positions of the next objects (sofa and armchair) are then defined with respect to the shelf's sides, clearly *not* based on *gaze* perspective, though this is not made explicit. Other descriptions furthermore employ the direction-neutral lateral term *neben* (beside), which also points to an embedded perspective: using the speaker's current point of view, the lateral side of the objects would denote a different location from the one described.

We analyzed the perspective chosen specifically for the four objects just mentioned (near or at the lefthand wall), which were the only ones with clear intrinsic fronts that were facing laterally into the room. Altogether 52.0% of all participants (21 in the functional and 31 in the nonfunctional condition) referred to at least one of these objects using an embedded kind of perspective (mental tour or object-based). This may reflect a desire to avoid confusion with the objects' intrinsic fronts when using *in front of* and *behind*, which express the objects' locations with the *gaze* perspective. This observation is paralleled by a similar effect with respect to the other side of the arrangement, where an embedded perspective yields linguistic expressions that are (in this particular setting) indistinguishable from the *gaze*-based (external) perspective with a *tandem* (rather than the more usual *mirror*) interpretation (e.g., describing the table as being *in front of the sofa* in [Figure 1](#)). Altogether we found 17 possible tandem interpretations for object descriptions; these were produced by 14

different speakers (with none producing more than two such utterances). 14 of the 17 descriptions occurred in the functional condition; 12 of these referred to the group of objects on the right side of the arrangement ([Figure 1](#)). This provides further indication of how the specific features of the configuration – namely, the objects' position and orientation within the arrangement – affected the speakers' choices.

In summary, speakers predominantly used an external gaze perspective (their own actual view on the scene rather than any imaginary movement), although other options were also available and indeed used by some speakers and with respect to specific kinds of configurations. In particular, the parts of the arrangement that supported embedded perspectives were those where objects' intrinsic fronts and back sides (rather than their lateral axis) could be used for reference.

**Choice of relatum.** Almost all speakers who described object locations used a relatum to do so, as did most speakers who described the orientation of an object. Overall, 95.6% of all units containing location information (produced by all 100 speakers) and 76.7% of all units containing orientation information (produced by 67 speakers) referred to an identifiable relatum.

We then asked which types of relatum speakers chose out of the available range of options. For this purpose, we analyzed these descriptions further with respect to the distribution of *another object*, *the environment*, and *the speaker* as relatum for descriptions of location in contrast to descriptions of orientation. Speakers produced different relatum type descriptions to varying degrees when referring to the location vs. the orientation of objects, displayed in [Figure 3](#).

--- Insert Figure 3 about here ---

Although the preferences for a particular type of relatum for locational descriptions did not generally coincide with the preferences for orientational descriptions, the encompassing environment and its features were prominent choices that were mentioned to the same extent in both location and orientation descriptions. The following units are typical examples taken from our data, with (4) exemplifying locational and (5) orientational descriptions:

- (4) This sofa is uhm located at the rearmost wall.

[*German original*: dieses Sofa steht ähm an der hintersten Wand]

- (5) The back points to the left side.

[*German original*: der Rücken der zeigt zur linken Seite]

Both types of information can be combined in units like the following:

- (6) The first chair stands with its back to the rear wall.

[*German original*: der erste Stuhl der steht mit dem Rücken zur hinteren Wand]

Locational descriptions rarely used the speaker's position as a relatum (16 instances produced by 14 different speakers). In 10 of these cases they did not refer to a single object but rather to the whole array – or a section of it – as locatum, as in the following example:

- (7) So now comes the front wall that lies directly in front of me.

[*German original*: so jetzt kommt die Vorderwand die direkt vor mir liegt]

In contrast, orientational descriptions regularly used the speaker as relatum, as in:

- (8) The chair stands with its seat directed towards me.

[*German original*: der Stuhl steht mit der Sitzseite zu mir gewandt]

These differences between choices of relatum for descriptions of location vs. orientation were statistically reliable, as shown by a 2 (unit types: location versus orientation) x 4 (relatum type: object as relatum, environment as relatum, speaker as

relatum, other relata) x 2 (array type: functional versus non-functional) mixed ANOVA on the mean percentages produced across participants. There was a significant interaction between unit type and relatum type,  $F(3, 294)=73.33$ ,  $p<0.0001$ . Thus, speakers used their own position more frequently in describing the orientation of objects than the location of objects, ( $p<0.0001$ ). Instead, for location descriptions they used *another object* and *other* relata more frequently than with orientation descriptions, (both  $p<0.01$ ).

There was also a marginally significant interaction between unit type and array,  $F(1, 98)=3.58$ ,  $p=0.06$ , which can be traced back to the greater percentage use of both location and orientation content in the non-functional compared to the functional array mentioned above (Table 1). None of the other effects or interactions were reliable.

Next, we asked whether speakers preferred to use objects which had just been mentioned to describe the position of new objects. In the case of locational units, we found that this was overwhelmingly the case. Of the 642 units (produced by 99 speakers) that contained an object as relatum of location, 97.2% relied on an object that had just been mentioned (in one of the three previous units). In only 18 cases (produced by 15 speakers) did they refer to objects mentioned earlier than that. Four of these cases actually contained an explicit marking of this fact (and others had extended object descriptions to support the reference), as in:

- (9) In fact, it stands, so to speak, to the right beside the yellow table that I described beside the big table that stands there.

[*German original*: und zwar steht das dann quasi rechts neben dem gelben Tisch den ich beschrieben hab' neben dem großen Tisch der da steht]

In contrast, when describing the *orientation* of an object (i.e., the direction in which the currently described object is pointing) by referring to another object as relatum,

speakers did not seem to have an equally strong preference for objects that had just been mentioned in the previous discourse. Of the 86 units (produced by 41 speakers) that contained an object as relatum of orientation, 28 (32.6 %; produced by 33 speakers) used an object that had not been mentioned in one of the three previous units.

## Discussion

How do speakers organize spatial concepts in order to linguistically convey information about a complex table-top configuration? We addressed this question in an analysis of natural descriptions of spatial arrangements collected from 100 speakers, focusing on their choices of description content, relatum, and perspective. Our results show that speakers draw on a large repertory of options with respect to linguistic choices and description strategies, exhibiting a number of systematic principles and patterns for spatial discourse. In other words, we can see from our data what speakers usually do, but also what they sometimes do and what they do not do. In summary, the speakers in our study described objects in a complex array overwhelmingly using their current *perspective* on the scene, with systematic exceptions to this rule based on features of the object arrangement. Location descriptions were based predominantly on *relative reference frames*, and they frequently employed the encompassing *environment as a relatum*. Orientation descriptions, in contrast, regularly related to the *speaker* in order to provide a direction. In the following, we discuss each of these findings in turn.

**Perspective.** Descriptions were overwhelmingly consistent with the external *gaze-based perspective*, which corresponds to the actual view speakers have on the scene when describing. Thus, the descriptions seldom depended on an imagined point of view within the scenario, or on object-based perspectives. This finding corresponds

with theories that base human location concepts predominantly on perceptual access and embodiment (e.g., Grabowski, 1999; Mallot & Basten, 2009).

However, it contrasts with earlier findings by Linde and Labov (1975) who found that speakers imagined travelling through the scenario they were currently describing (the apartments they lived in), employing continually changing perspectives. The marked differences in perspective choice in our study and in theirs can be accounted for by two key differences in the way the studies were run. First, the scales and configurations were very different; the arrays we used were in peripersonal (near) space and involved equidistant objects, while the Linde and Labov study involved extrapersonal (far) space with naturally clustered object environments. Our results thus support previous evidence indicating that spatial language use is constrained by scale (Lautenschuetz et al., 2007) and configuration (Ehrich & Koster, 1983). Second, the participants in our study saw the visual scene concurrently with description while Linde and Labov had participants describe from memory. These two differences combined may have led participants to simulate walking around their apartments in Linde and Labov's task, consistent with data showing that distance between rooms is monotonically related to the time taken to make judgements about the relative positions of rooms in a memory paradigm (e.g., Rinck & Denis, 2004).

So the scale and configuration of a given scenario and the discourse task (description or memory) may have a substantial influence on the overall spatial description strategy, even in the absence of actual navigation (true for both scenarios). A route perspective may be encouraged if one imagines going through a doorway and walking through the scene using typical paths; in contrast, the instruction to describe an arrangement such that another person could re-arrange the items may have

encouraged a small-scale conceptualization corresponding to the actual view on the scene.

In a study by Ehrich (1985) gaze perspective was associated with orderly sequences of local object description, while the object-based (embedded) perspective tended to yield more irregular, discontinuous structures in the descriptions. This finding corresponds to the idea expressed by Tversky et al. (1999) that orienting towards intrinsic object features requires jumping between objects instead of proceeding sequentially. However, the present study does not support this assumed correspondence between embedded perspective choices and discontinuous descriptions. While our data also contained distinct patterns of orderly and discontinuous sequences of individual object descriptions depending on the functional features of the configuration (see Andonova et al., in press), the current analysis nevertheless revealed a consistent general preference for the gaze-based perspective. Thus, contrary to the earlier assumptions based, for instance, on Ehrich's results (Ehrich, 1985), speakers may pursue orderly descriptions and switch perspectives locally at particular places; or they may stick to a consistent perspective while still jumping discontinuously between individual objects to be described.

Our findings highlight how particular features of the configuration can affect description strategies particularly with respect to the cases of embedded perspective usage. Both the orientation of objects within the scenario and the functional relationships between objects appeared to matter in this respect. This finding corresponds to Ehrich's (1985) earlier findings as well as to the substantial body of research investigating the effects of functional relationships between objects on the choice of spatial terms and reference frames (Coventry & Garrod, 2004; Carlson-Radvansky & Radvansky, 1996).

Perspective switches were not indicated linguistically, but had to be derived from the choice of spatial terms used for a particular configuration. Research in natural dialogue settings, enabling direct feedback, indicates the extent to which listeners are capable of reacting to cues in descriptions that indicate such a perspective switch (Schober 1993, 1995; Goschler, Andonova, & Ross, 2008). Further investigation of the dialogic processes involved here may highlight the means by which this is achieved, and reveal how and when communication may fail.

**Reference frames.** Levinson's (1996) seminal paper for the first time successfully disentangled various notions of reference frame by proposing a unified framework. In a nutshell, one may say that if the source for perspective (called *origin*) coincides with the relatum, the underlying conceptual reference frame is called *intrinsic*; if the source for perspective is distinct from the relatum, the reference frame is called *relative*. Prior to the introduction of this framework, the literature witnessed many different individual classifications of reference frames focusing variously on choice of perspective or on other kinds of relevant strategies, making generalization across studies difficult. For this reason a systematic account of speakers' preferences of reference frames in the context of complex arrangement descriptions has been absent. So how do the speakers' preferences for relatum and perspective combine in terms of intrinsic or relative reference frames?

Our speakers described the objects in the room either in relation to another object known from the previous discourse context – typically, the object directly described before – or in relation to the environment, with a slight preference for the former. Many of the local descriptions yield a valid interpretation with more than one possible reference frame, particularly if another object is used as relatum rather than the environment or the speaker. For instance, a chair positioned at the front wall, with

its back to the speaker, could be used as a relatum for another object placed on its left side either by using the speaker's current view on the scene or by assuming that the chair has an intrinsic left side. Further options for interpretation arise when imagining an embedded view as part of an imagined tour through the scene. Avoiding such conflations is virtually impossible in any configuration containing more than two objects, given that each pair of objects within the scene can be viewed and described from more than one (external or embedded) perspective.

For the interpretation of spatial descriptions, the conflation of several potentially valid reference frames does not pose any problems if the target location remains the same across interpretations. In fact, there is some evidence that speakers prefer spatial term assignments when they fit more than one reference frame (Carlson-Radvansky & Irwin, 1993). Thus, choosing a description that holds across various underlying reference frames may be a common and suitable communicative strategy. Because of these effects, assigning local reference frames may not always be feasible in complex configurations, and so we did not attempt this in the analysis framework adopted here.

Nevertheless, our perspective analysis still supports the conclusion that *relative* reference frames dominated over *intrinsic* ones in the present study. Most descriptions were consistent with the gaze-based perspective, which together with an object or the environment yields a relative reference frame. Since the environment (the "room", without a door or any other relevant features) did not possess any intrinsic sides, all descriptions of objects based on their position inside the room or relative to the room's walls, sides, or corners are necessarily gaze-based. Since this category covers more than one-third of the descriptions, it appears to be a reasonable strategy to stick to the same perspective also in those cases in which furniture items

rather than the environment served as a basis for reference. The cases of conflation of relative and intrinsic reference frames then supported the speakers in sticking to this perspective choice without conflict. Exceptions were triggered by particular configurations in which object functions (e.g., the relationship between armchairs and living room tables, or the interactionally prominent front sides of cupboards) apparently became relevant for a speaker. As a result, the number of clearly embedded perspectives (which include intrinsic reference frames as well as cases of potential mental rotation, which again correspond to relative reference frames) is overall very low. This finding contrasts with early suggestions that intrinsic reference systems may generally be predominant (Miller & Johnson-Laird, 1976, in part supported by Ehrich, 1985) but replicates Levinson's (2003) more recent finding that, in Western cultures, relative reference systems are generally preferred (although specific preference is situationally constrained).

In general, two main factors appear to interact in inducing either relative or intrinsic frames. On the one hand, a complex scenario yields a tendency for stable gaze-based perspectives, leading to a consistent use of *relative* reference frames; on the other hand, clear intrinsic sides may lead to an increased use of *intrinsic* object-based reference frames. Further research should directly address this apparent trade-off by varying the obviousness of the objects' intrinsic sides, their arrangement in functional ways, and the complexity of the configuration to be described. Furthermore, adding intrinsic features to the room itself, such as a door, may further decrease speakers' adherence to a consistent gaze-based strategy, which might result in an increase in intrinsic reference frame use.

**Environment-based relata.** The option of using the encompassing environment (the "room" and its features) as a relatum in the present scenario

apparently plays a major role for complex configuration descriptions, suggesting that the global structure of a complex but clearly delimited configuration is a salient aspect of the spatial environment. This finding is consistent with Shanon's (1984) study investigating actual room descriptions, which however did not address reference systems. Carroll and von Stutterheim (1993) found frequent occurrences of global spatial structures, but their analysis focuses on differences between English and German in referring to encompassing (environment-based) relata. It is mostly in Talmy's work (e.g., Talmy, 2000) that frequent consideration of environment-based relata can be found, associated with his much cited observations concerning the *ground* (relatum) as being systematically larger, more salient, and less movable than the *figure* (locatum). Nevertheless, these findings are seldom accounted for in discussions of reference frames, which almost exclusively focus on *external* relationships between objects (but see Tenbrink, in press).

**Orientation.** How speakers choose the relatum of an *orientation* description (in cases where a relatum is used in such a description at all; a relatum is not obligatory as a direction is sufficient) has hitherto evaded empirical investigation. Although choice of relatum for orientation description might theoretically have simply mirrored the choice of relatum in a *location* description, this was not the case in our data. While location descriptions typically either re-used the object just described (i.e., the locatum of a previous location description) or the global structure as relatum, orientation descriptions relied more heavily on the global structure but also used the *speaker* as a relatum quite frequently, using descriptions such as *the chair points to me*. This option was almost absent from the location descriptions, for the obvious reason that all objects were located in front of the speaker – a fact which, in turn, did not affect the feasibility of using this kind of relatum for orientation

descriptions. Thus, the description of how objects are *oriented* is based on different principles than the description of where they are *located*.

## Conclusion

Most investigations of spatial descriptions have focused on the relationship between only two objects at a time. Our study addressed how speakers describe complex object arrangements placed directly in front of them, focusing on the type of information provided, the nature of the relatum, and the underlying perspective chosen for spatial description. The analysis revealed a consistent default strategy of choosing the external gaze-based perspective, which could be overruled locally based on particular features of the configuration. This global perspective choice came along with systematic preferences for entities chosen as relatum, including the environment as a whole (and its features), as well as objects mentioned in directly preceding units. Contrary to other kinds of spatial scenarios, objects' locations were typically not described in relation to the speaker's current or imagined position; however, this position regularly served as a basis for the specification of the *orientation* of objects.

Our results open up a range of avenues for further research. First of all, the interplay between general discourse strategies – such as adhering to a particular perspective – and local configuration features clearly warrants further investigation. In particular, the conceptual transition from a local object-to-object relation to a more complex scene with further objects influencing individual object descriptions is currently not well understood. Second, further investigation of orientation information accompanying location information may highlight how speakers' strategies for conveying spatial details interact with the availability of potential relata, as well as with global description strategies chosen for a particular setting. Third, discourse strategies such as those identified here may immediately change as soon as an

addressee is directly involved; thus, the investigation of spatial dialogue about complex arrangements is a fruitful avenue for future research. Finally, these strategies and principles are certainly mediated by individual differences based on personal experience, levels of spatial ability, habits and proclivities, and cultural factors, all of which are subject to much ongoing research in spatial cognition (e.g., Hegarty et al., 2006).

## Acknowledgements

We gratefully acknowledge funding for the present work by the DFG for the SFB-TR 8 Spatial Cognition Bremen / Freiburg, project I5-[DiaSpace], from the Hanse-Institute in Delmenhorst (HWK) for a fellowship awarded to Kenny Coventry, and from the Volkswagen Foundation for a tandem project awarded to Jan Wiener and Thora Tenbrink. We thank our student assistants for the many hours of data collection, transcription, and annotation that went into this study.

## References

- Andonova, E., Tenbrink, T., & Coventry, K.R. (in press). Function and context affect spatial information packaging at multiple levels. *Psychonomic Bulletin & Review*.
- Bartlett, F. C. (1932). *Remembering: A case study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Bateman, J., Tenbrink, T., & Farrar, S. (2007). The role of conceptual and linguistic ontologies in discourse. *Dialogue modelling: Computational and empirical approaches. Special Issue of Discourse Processes*, 44(3), 175–213.

- Berti, A., & Rizzolatti, G. (2002). Coding near and far space. In H. –O. Karnath, A. D. Milner, & G. Vallar (Eds.), *The cognitive and neural bases of spatial neglect* (pp. 119-129). New York: Oxford University Press.
- Carlson-Radvansky, L. A., & Irwin, D.E. (1993). Frames of reference in vision and language: Where is above? *Cognition*, 46, 223-244.
- Carlson-Radvansky, L. A., & Radvansky, G. A. (1996). The influence of functional relations on spatial term selection. *Psychological Science*, 7, 56-60.
- Carroll, M., & von Stutterheim, C. (1993). The representation of spatial configurations in English and German and the grammatical structure of locative and anaphoric expressions. *Linguistics* 31, 1011-1041.
- Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.
- Coventry, K. R., & Garrod, S. C. (2004). *Saying, seeing and acting: The psychological semantics of spatial prepositions*. Hove and New York: Psychology Press.
- Coventry, K. R., Prat-Sala, M., & Richards, L. (2001). The interplay between geometry and function in the comprehension of ‘over’, ‘under’, ‘above’ and ‘below’. *Journal of Memory and Language*, 44, 376-398.
- Coventry, K. R., Valdés, B., Castillo, A., & Guijarro-Fuentes, P. (2008). Language within your reach. Near-far perceptual space and spatial demonstratives. *Cognition*, 108, 889-895.
- Denis, M. (1997). The description of routes: A cognitive approach to the production of spatial discourse. *Cahiers de Psychologie Cognitive*, 16(4), 409-458.
- Denis, M., Pazzaglia, F., Cornoldi, C., & Bertolo, L. (1999). Spatial discourse and navigation: an analysis of route directions in the city of Venice. *Applied Cognitive Psychology*, 13(2), 145-174.

- de Vega, M., Rodrigo, M. J., Ato, M., Dehn, D. M., & Barquero, B. (2002). How nouns and prepositions fit together: An exploration of the semantics of locative sentences. *Discourse Processes*, 34(2), 117-143.
- Ehrich, V. (1985). Zur Linguistik und Psycholinguistik der sekundären Raumdeixis [On the linguistics and psycholinguistics of secondary spatial deixis.] In H. Schweizer (Ed.), *Sprache und Raum: Ein Arbeitsbuch für das Lehren von Forschung* (pp. 130-161). Stuttgart: Metzler.
- Ehrich, V., & Koster, C. (1983). Discourse organization and sentence form: The structure of room descriptions in Dutch. *Discourse Processes*, 6, 169-195.
- Goschler, J., Andonova, E., & Ross, R. (2008). Perspective use and perspective shift in spatial dialogue. In C. Freksa, N. Newcombe, P. Gärdenfors, & S. Wölfl (Eds.), *Spatial Cognition VI: Learning, reasoning, and talking about space* (pp. 250-265). Berlin: Springer.
- Grabowski, J. (1999). A uniform anthropomorphological approach to the human conception of dimensional relations. *Spatial Cognition and Computation*, 1, 349-363.
- Grabowski, J., & Weiss, P. (1996). The prepositional inventory of languages: a factor that affects comprehension of spatial prepositions. *Language Sciences*, 18, 19-35.
- Hegarty, M., Montello, D.R., Richardson, A.E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, 34, 151– 176.
- Herrmann, T. (1990). Vor, hinter, rechts und links: das 6H-Modell. Psychologische Studien zum sprachlichen Lokalisieren [In-front, behind, right and left: the 6H model. Psychological studies of verbal localization]. *Zeitschrift für Literaturwissenschaft und Linguistik*, 78, 117-140.

- Herrmann, T., & Grabowski, J. (1994). *Sprechen. Psychologie der Sprachproduktion*. Heidelberg: Spektrum.
- Herskovits, A. (1986). *Language and spatial cognition*. Cambridge: Cambridge University Press.
- Hill, C. (1982). Up/down, front/back, left/right. A contrastive study of Hausa and English. In J. Weissenborn & W. Klein (Eds.), *Here and there. Cross-linguistic studies on deixis and demonstration* (pp. 13-42). Amsterdam: John Benjamins.
- Hirtle, S. C., & Mascolo, M. F. (1986). The effect of semantic clustering on the memory of spatial locations. *Journal of Experimental Psychology: Learning, Memory and Cognition, 12*, 181-189.
- Klabunde, R. (1999). Logic-based choice of projective terms. In: W. Burgard, T. Christaller, A.B. Cremers (Eds.), *KI-99, LNAI 1701* (pp. 149-158). Berlin, Heidelberg: Springer.
- Klein, W. (1979). Wegauskünfte [Route descriptions]. *Zeitschrift für Literaturwissenschaft und Linguistik (LiLi), 33*, 9-57.
- Kriz, S., & Hegarty, M. (2005). Spatial perspective in spoken descriptions of real world environments at different scales. In B.G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the XXVII annual conference of the Cognitive Science Society* (pp. 1196-1201). Mahwah, NJ: Lawrence Erlbaum Associates.
- Làdavas, E. (2002). Functional and dynamic properties of visual peripersonal space. *Trends in Cognitive Science, 6*, 17-22.
- Lautenschuetz, A. -K., Davies, C., Raubal, M., Schwering, A., & Pederson, E. (2007). The influence of scale, context and spatial preposition in linguistic topology. In: T. Barkowsky, M. Knauff, G. Ligozat, & D. Montello (Eds.), *Spatial Cognition V: Reasoning, action, interaction* (pp. 439-452). Berlin: Springer.

- Levelt, W. J. M. (1996). Perspective taking and ellipsis in spatial descriptions. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and space* (pp. 77-107). Cambridge, MA: MIT Press.
- Levinson, S. C. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and space* (pp. 109-169). Cambridge, MA: MIT Press.
- Levinson, S. C. (2003). *Space in language and cognition*. Cambridge: Cambridge University Press.
- Linde, C., & Labov, W. (1975). Spatial networks as a site for the study of language and thought. *Language*, 51, 924-939.
- Mainwaring, S. D., Tversky, B., Ohgishi, M., & Schiano, D. J. (2003). Descriptions of simple spatial scenes in English and Japanese. *Spatial Cognition and Computation* 3(1), 3-42.
- Mallot, H. A., & Basten, K. (2009). Embodied spatial cognition: Biological and artificial systems. *Image and Vision Computing* 27:1658–1670.
- McNamara, T. P., & Diwadkar, V. A. (1997). Symmetry and asymmetry of human spatial memory. *Cognitive Psychology*, 34, 160-190.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge: Cambridge University Press. Formatted: French (France)
- Montello, D. (1993). Scale and multiple psychologies of space. In A. U. Frank & I. Campari (Eds.), *Spatial Information Theory: A theoretical basic for GIS*. (pp.312-321). Berlin: Springer. Lectures Notes in Computer Science.
- Newcombe, N., Huttenlocher, J., Sandberg, E., Lie, E., & Johnson, S. (1999). What do misestimation and asymmetries in spatial judgement indicate about spatial

- representation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, 986-996.
- Pederson, E. (2003). How many reference frames? In: C. Freksa, W. Brauer, C. Habel, & K. F. Wender (Eds.), *Spatial Cognition III: Routes and navigation, human memory and learning, spatial representation and spatial learning* (pp. 287-304). Berlin: Springer.
- Plumert, J.M., Carswell, C., de Vet, K., & Ihrig, D. (1995). The content and organization of communication about object locations. *Journal of Memory and Language*, 34, 477-498.
- Plumert, J. M., Spalding, T. L., & Nichols-Whitehead, P. (2001). Preferences for ascending and descending hierarchical organization in spatial communication. *Memory & Cognition*, 29, 274-284.
- Rinck, M., & Denis, M. (2004). The metrics of spatial distance traversed through mental imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 1211-1218.
- Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language* 50, 696-735.
- Sadalla, E.K., Burroughs, W.J., & Staplin, L.J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 516-528.
- Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition* 47:1-24.
- Schober, M.F. (1995). Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about location? *Discourse Processes*, 20(2), 219-247.
- Selting, M. (2000). The construction of units in conversational talk. *Language in Society*, 29, 477-517.

- Shanon, B. (1984). Room descriptions. *Discourse Processes*, 7, 225-255.
- Talmy, L. (2000). *Toward a cognitive semantics*, 2 vols. Cambridge, MA: MIT Press.
- Taylor, H. A., Naylor, S. J., Faust, R. R., & Holcomb, P. J. (1999). "Could you hand me those keys on the right?" Disentangling spatial reference frames using different methodologies. *Spatial Cognition and Computation*, 1(4), 381-397.
- Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language*, 35, 371-391.
- Tenbrink, T. (in press). Reference frames of space and time in language. *Journal of Pragmatics*.
- Tversky, B. (1999). Spatial perspective in descriptions. In P. Bloom, M.A. Peterson, L. Nadel, & M.F. Garrett (Eds.), *Language and space* (pp. 109-169). Cambridge, MA: MIT Press.
- Tversky, B., & Hard, B.M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, 110, 124–129.
- Tversky, B., Lee, P., & Mainwaring, S. (1999). Why do speakers mix perspectives? *Spatial Cognition and Computation*, 1(4), 399-412.
- Tyler, A., & Evans, V. (2003). *The semantics of English prepositions: Spatial sciences, embodied meaning, and cognition*. Cambridge: Cambridge University Press.
- Ullmer-Ehrich, V. (1979). Wohnraumbeschreibungen [Living space descriptions]. In: *LiLi, Zs. für Literaturwissenschaft und Linguistik* 33, 58-83.
- Vorwerg, C. (2009). Consistency in successive spatial utterances. In K. R. Coventry, T. Tenbrink, & J. Bateman (Eds.), *Spatial language and dialogue* (pp. 40-55). Oxford: Oxford University Press.

- Watson, M. E., Pickering, M. J., & Branigan, H. P. (2004). Alignment of reference frames in dialogue. In K.D. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26<sup>th</sup> annual conference of the Cognitive Science Society* (pp. 1434-1440). Hillsdale, NJ: Lawrence Erlbaum.
- Weiss, P., Grabowski, J., & Miller, G. A. (1996). Factors affecting spatial deictic communication: A comparison of German and American English. In *Proceedings of the second colloquium on deixis "Time, space and identity"*, March 28-30 1996, Nancy (pp. 45-73). Nancy: Centre de Recherche en Informatique.
- Zwarts, J. (2003). Vectors across spatial domains: From place to size, orientation, shape, and parts. In E. van der Zee & J. Slack (Eds.), *Representing direction in language and space* (pp. 39-68). Oxford: Oxford University Press.

Table 1: Mean percent utterances of different description content by array.

	Functional array	Non-functional array
Location information only	74.0	62.6
Orientation information only	4.7	6.0
Both location & orientation	10.8	22.0
No location or orientation	10.5	9.4
Total	100.0	100.0

Table 2. Annotation examples (Example 1 and 2). LD = locational description; NLO = neither locational nor orientational information; CG = compatible with gaze; PN = perspective neural; E = environment; KO = known object; O = other

<b>Example 1, Functional Arrangement</b>	<b>Content</b>	<b>Per-spective</b>	<b>Relatum of Loca-tion</b>
also vorne in der Mitte steht ein Tisch ein gelber Tisch (okay in the front in the middle there is a table a yellow table)	LD	CG	E
links und rechts davon zwei Stühle (left and right of it two chairs)	LD	CG	KO
weiter links ganz in der Ecke ein kleines niedriges Regal (further left all the way in the corner a small low cupboard)	LD	CG	E
ganz rechts in der Ecke ein Sofa mit einem Kissen drauf (all the way to the right in the corner there is a sofa with a cushion on it)	LD	CG	E
oben links steht eine Vitrine (in the top left there is a cabinet)	LD	CG	E
zwischen der Vitrine und dem kleinen Regal links an der Wand ein hohes Regal (between the cabinet and the small cupboard on the left at the wall there is a high shelf)	LD	CG	O
rechts in der Ecke stehen zwei Sessel ein Stückchen auseinander gestellt mit zwei Kissen drauf (in the right corner there are two armchairs placed at some distance from each other with two cushions)	LD	CG	E
weiter links davon in der Mitte oben steht ein kleinerer Sessel (further to the left of that in the middle top there is a smaller armchair)	LD	CG	O
ganz äh weiter links davon noch nicht ganz in der Ecke steht ein weiteres Regal (all the way uh further to the left of that, not quite in the corner yet there is a further cupboard)	LD	CG	O
ganz rechts ähm ein Stückchen abgerückt von Rand steht ein ähm niedrigerer gelber Tisch (all the way to the right uh a bit apart from the edge there is a lower yellow table)	LD	CG	E
und ganz in der Mitte stehen zwei Hocker (and directly in the middle there are two stools)	LD	PN	E
und links weiter ein Stuhl (and further to the left a chair)	LD	CG	E
<b>Example 2, Nonfunctional Arrangement</b>			<b>Relatum of Loca-tion</b>
ja also oben links in der Ecke ist ein Sofa mit einem Kissen drauf (in the top left corner there is a sofa with a cushion)	LD	CG	E
davor also unten drunter im Prinzip ein Regal (in front of that or below it there is a shelf)	LD	CG	KO
darunter links in der Ecke ein ja wie nennt man das Ding nochmal Sessel mit Kissen drauf (below that in the left	LD	CG	O

<i>corner there is a, how do you call this thing again, an armchair with a cushion on it)</i>			
dann oben rechts neben dem Sofa hat man einen Stuhl ( <i>then on the top right beside the sofa there's a chair</i> )	LD	CG	O
dadrunter ungefähr auf Höhe des Schrankes noch ein Stuhl ( <i>below that, roughly at the level of the shelf, there's another chair</i> )	LD	CG	O
und dann unten auf Höhe des Sessels einen kleinen Tisch halbrund nee viertelrund ( <i>and then at the bottom, at the level of the armchair, a small table, a half-round, no, quarter of a circle</i> )	LD	CG	O
dann oben rechts wieder neben dem Stuhl eine Art Sessel ( <i>Then on the top right again beside the chair there's a kind of armchair</i> )	LD	CG	O
ist ein komisches Couchelement oder so ( <i>that's a strange couch element or so</i> )	NLO	PN	n.a.
dadrunter wieder ein Stuhl und unten ein Tisch ( <i>below that there is again a chair and on the bottom a table</i> )	LD	CG	O
nochmal oben rechts ein Regal so ein Wandregal mit drei Ebenen ( <i>again, on the top right, there's a shelf, a kind of wall shelf with three levels</i> )	LD	CG	E
dadrunter so einen viertelrunden Tisch ( <i>below that another quarter-circle table</i> )	LD	CG	KO
und ganz unten ein kleines Regal mit zwei Ebenen ( <i>and all the way down a small shelf with two levels</i> )	LD	CG	E
oben rechts in der Ecke hat man ein Sessel mit einem Kissen drauf ( <i>in the top right corner there is an armchair with a cushion</i> )	LD	CG	E
in der Mitte einen Tisch und unten eine Anrichte ( <i>in the middle there's a table and on the bottom a sideboard</i> )	LD	CG	O
und das ist ganz rechts in der Ecke ( <i>and that is all the way to the right in the corner</i> )	LD	CG	E

### **Figure captions**

#### **Figure 1**

Functional array. Left: View from above. Right: View from the observer as used in the study. The objects are (left to right, top to bottom): row 1: cupboard, side table with two wheels, fold-out sofa, armchair with yellow cushion, armchair with blue cushion; row 2: cupboard with three shelves, chair, two small corner tables, low living room table; row 3: corner unit, chair, dinner table, chair, sofa with yellow cushion.

#### **Figure 2**

Non-functional array. Left: View from above. Right: View from the observer as used in the study. The objects are (left to right, top to bottom): row 1: sofa with yellow cushion, chair, fold-out sofa, corner unit, armchair with blue cushion; row 2: cupboard with three shelves, two chairs, small corner table, low living room table; row 3: armchair with yellow cushion, small corner table, dinner table, side table with two wheels, cupboard.

#### **Figure 3**

Mean percent utterances of different relata in descriptions of location and orientation.

Figure 1

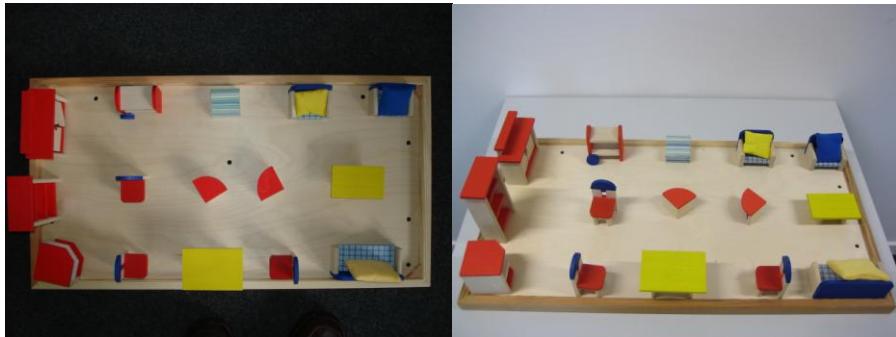


Figure 2

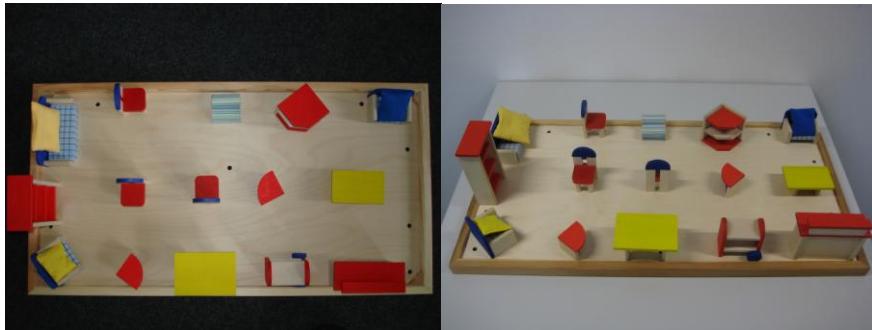


Figure 3

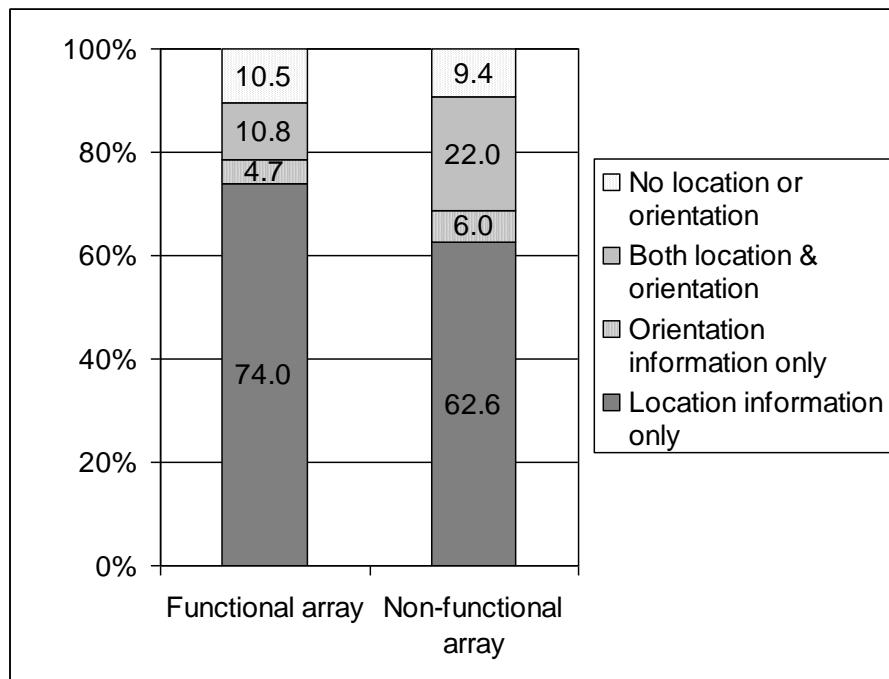


Figure 4

