Semantics and Application of Spatial Dimensional Terms in English and German

Thora Tenbrink

SFB/TR 8 Report No. 004-03/2005
Report Series of the Transregional Collaborative Research Center SFB/TR 8 Spatial Cognition
Universität Bremen / Universität Freiburg

Contact Address:

Dr. Thomas Barkowsky
SFB/TR 8
Universität Bremen
P.O.Box 330 440
28334 Bremen, Germany

Tel +49-421-218-8625
Fax +49-421-218-8620
barkowsky@sfbtr8.uni-bremen.de
www.sfbtr8.uni-bremen.de

> "Spatial language is used primarily to indicate where things are; indicating where things are is frequently an important aspect of identifying or referring to them; identification and reference are critical aspects of linguistic communication."
>
> (Miller & Johnson-Laird 1976:410)

Semantics and application of spatial dimensional terms in English and German

Thora Tenbrink, Faculty of Linguistics and Literary Sciences, University of Bremen

This report presents a summary and discussion of findings in the literature with respect to the linguistic repertory (syntactic variability), semantics, and application of spatial dimensional terms in English and German, such as *left/links* or *above/oben*. These terms are called dimensional because they describe the relation of two entities to each other on a specific spatial dimension. Their application depends on a wide range of situation- and discourse-related factors, such as the conceptualisation of underlying reference systems, the nature of the entities to be related, and the discourse task. Each of these issues will be dealt with separately in order to determine their specific impact on the interpretation and applicability of the spatial terms.

I will start with a detailed presentation of the range of variety in spatial language, with specific focus on dimensional terms, and then turn to the wide range of aspects coming into play in their application. Many important findings are not restricted to dimensional terms, but it is only with respect to dimensional terms that I aim towards completion and detail in presentation to the greatest possible extent.

A number of interesting insights on the variability, features, and applicability of English spatial terms is collected in the seminal work of Herskovits (1986). I will start from her insights and supplement these with more recent findings as well as (contrasting or equivalent) observations pertaining to the German language, relying on a number of sources. Most literature on spatial terms presents syntactic variability and semantic aspects of the various expressions along with various pragmatic factors influencing their application. My aim in the present work is to keep these issues apart as far as possible, in order to present, on the one hand, an overview of the linguistic repertory, and on the other, an account of the range of discourse-related and discourse-external factors coming into play in the application of this repertory. Although this is clearly a simplification in light of the fact that language is only meaningful in its application in discourse, I will start from Herskovits' assumption that there is a semantic core for every spatial expression. This assumption will be discussed and modified in the first section of the present work. The following sections then examine findings with regard to pragmatic factors, including considerations of underlying reference systems (starting from the systematics provided by Levinson 2003), and functional factors influencing applicability (such as those addressed in the work of Carlson, Coventry, Garrod, and related research).

1 Linguistic repertory: Syntactic variability and semantic cores of spatial terms

1.1 Categories of spatial expressions

To begin with, a clarification of terminology and categorisation is in order. In the present work, following Wunderlich & Herweg (1991:778), I use the category name **dimensional** for the terms under analysis:

> "Eine geschlossene Klasse von in der Regel sechs Ausdrücken bilden die dimensionalen Präpositionen: *vor, hinter, über, unter, rechts, links.*"

> [A closed class of usually six expressions is constituted by the dimensional prepositions *vor (in front of), hinter (behind), über (above), unter (below), rechts (right), links (left)*.]

This category name highlights the fact that the terms in this category are used to express relative location on the spatial dimensions or axes. In the present work, the name **dimensional** is used for all syntactic forms that the terms can take in natural language, i.e., not only prepositions. The same category name applies likewise and unproblematically to related expressions in the temporal domain, unlike the near-synonymous term **projective** which is also used widely in the literature (e.g., Herskovits 1986, Eschenbach 2004) for the same class of spatial expressions. Another term that is sometimes used for spatial terms in a similar way, for example, in van der Zee & Slack (2003) is **directional**; however, this term is highly polysemous, comprising orientation (of an entity, i.e., object or person), location in a specific area with regard to another entity, and direction of movement. Indeed, elsewhere in the literature the term is usually associated with a predominantly **dynamic** interpretation that is not in focus in the present work, since motion involves both spatial and temporal dimensions which are kept apart as far as possible here. For instance, Wunderlich & Herweg (1991:759) distinguish **Lokale** and **Direktionale** (similarly: **locative** vs. **directional** in Eschenbach, 2004) in the realm of spatial expressions:

> "Als 'Lokale' sollen in diesem Artikel Ausdrücke einer Sprache bezeichnet werden, die zur Lokalisierung von Objekten oder Ereignissen dienen. Die Variable 'Zeit' findet dabei nur insoweit Eingang, als sie im Tempus des Verbs oder im Zeitadverbial kodiert ist, wodurch die Lokalisierungssituation zeitlich eingeordnet wird. Als 'Direktionale' sollen Ausdrücke bezeichnet werden, die die Veränderung der Lokalisierung eines Objektes (also einen Ortswechsel) ausdrücken (…) Die Variable 'Zeit' findet direkt Eingang in der Kodierung des Ortswechsels."
>
> [In this article, 'Lokale' (locatives) denote linguistic expressions that serve to localise objects or events. The variable 'time' only participates by being encoded in the tense of the verb or in a temporal adverbial, providing a temporal time frame for the localising situation. 'Direktionale' (directionals) denote expressions that represent changes in the localisation of an object (i.e., a change of place). (…) The variable 'time' is directly represented in the encoding of a change of place.]

Klein (1991:89) points to the fact that the so-called directionals cannot really express a direction since this would require information about both a source and a destination:

> "Der gängige Ausdruck 'Direktionale' für Raumausdrücke wie 'in den Garten' oder 'nach Heidelberg' ist bestenfalls irreführend. Solche Ausdrücke charakterisieren keine Richtung. Es gibt beliebig viele Richtungen mit dem gemeinsamen Zielort 'Garten' oder 'Heidelberg'. Um eine Richtung zu charakterisieren, sind zunächst einmal zwei Ortsangaben erforderlich (von denen eine möglicherweise dem Kontext zu entnehmen ist)."
>
> [The prevalent expression 'directionals' for spatial expressions such as 'in den Garten' (into the garden) or 'nach Heidelberg' (to Heidelberg) is at best misleading. Such expressions do not characterise a direction. There are arbitrarily many directions that share the destination 'garden' or 'Heidelberg'. In order to characterise a direction, to begin with, two specifications of locations are required (one of which may possibly be inferred from the context).]

He proposes to use the expression **Destinative** instead. However, Klein's proposal has not been taken up in the literature, probably because the source of direction is usually in fact provided by the context so that this distinction turns out to be irrelevant. Viewed another way, the only true directionals would then be expressions that refer to absolute directions, such as *north, south, east, west* – but even these are generally employed in context to describe a direction from a specific (usually implied) starting point.

According to Miller & Johnson-Laird (1976:407), describing the location of moving objects builds on describing the location of stationary objects, and – in English – the full range of locative prepositions can be used for locating moving objects. Similarly, instead of describing the location of the path, its direction may be indicated, again using the same expressions as

for static locations. These overlaps may sometimes create ambiguities, reinforced by the fact that people do not always make use of available distinctions such as that between *in* and *into*, simply using *in* in colloquial speech.

Similarly, Herskovits (1986:8) points out that

> "static prepositions can be used in dynamic contexts (…) I exclude an in-depth study of dynamic prepositions, or of dynamic contexts for the static ones, but much of what holds true in the simple static cases carries over to dynamic contexts. Usually, some aspect of the action denoted by the verb involves a part of space, which is located with respect to the reference object by means of the preposition exactly as any static physical object would be."

Thus, it is assumed that an analysis of directional terms would yield a high degree of overlap with an analysis of static locative expressions (which is probably one reason why the term 'directional' is used for both static and dynamic situations in van der Zee & Slack 2003). Crucially, the verb contributes meaning to the overall construction that may influence the interpretation of the spatial term (Coventry & Garrod 2004a:10f.). In the present work, analysis starts from locative expressions; some specific (interpretational) observations with regard to directionals are presented in section 4 below. Several syntactic peculiarities of directionals are accounted for in the present section.

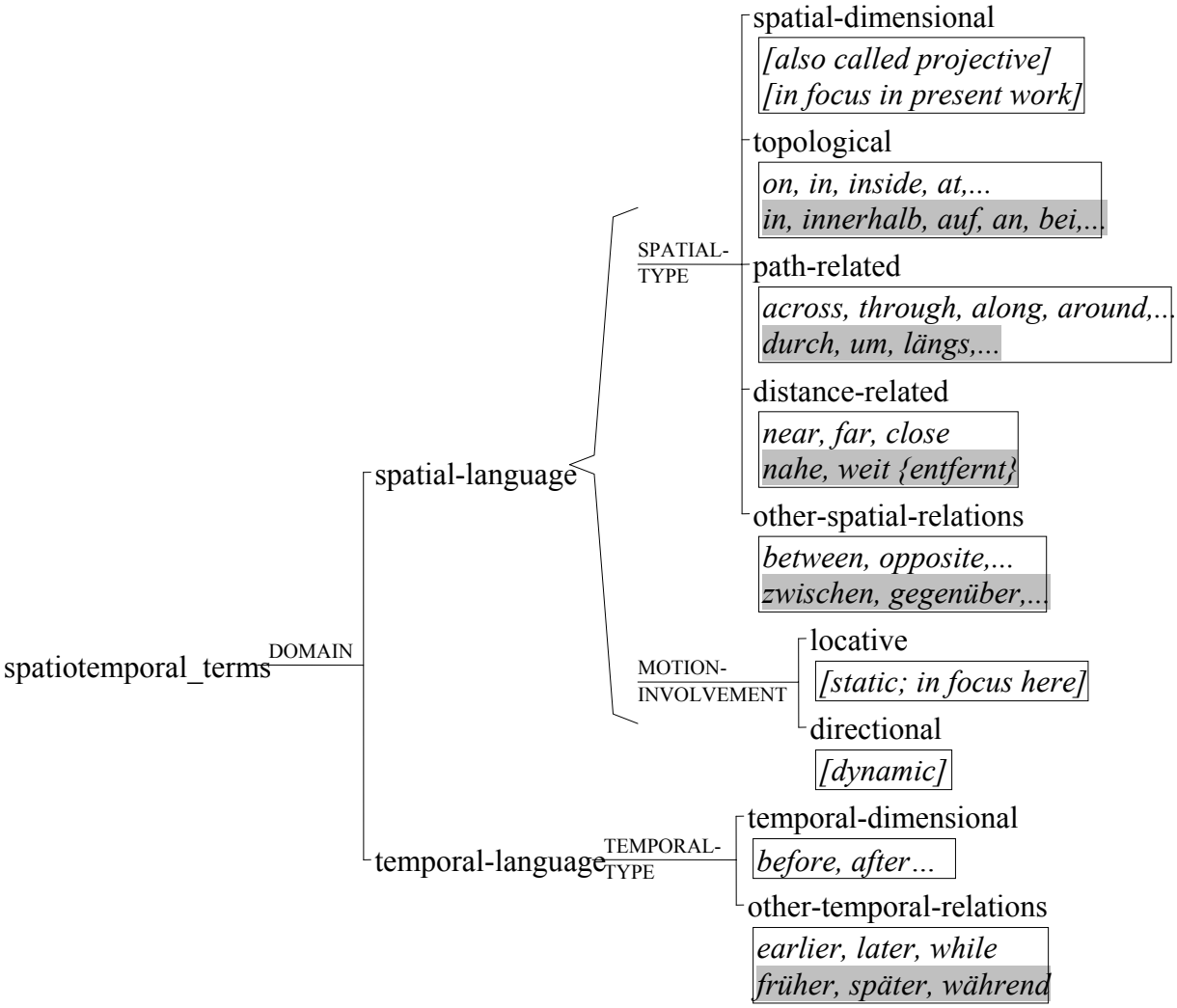


Figure 1. Overview of spatiotemporal expressions

The categories labeled 'locative' and 'directional' comprise diverse kinds of static and dynamic spatial terms (although these may sometimes be realized in different ways), i.e., also non-dimensional expressions such as *on/onto, in, across,* etc. Thus, dimensional expressions are contrasted from other classes of spatial expressions, e.g., **topological** (expressing

neighborhood or contiguity), **path-related**, **distance-related**, and other categories (see e.g. Pribbenow 1992 for a more detailed account). Furthermore, they should not be confused with **extensional** terms: these are sometimes (e.g., Lang 1989) also called 'dimensional' because they depend on the one-, two-, or three-dimensional extension of objects (cf. Clark 1973), e.g., *lang/kurz & long/short, dick/dünn & thick/thin*, etc. If all three dimensions are involved, one usually talks about *size* not extension, which concerns the application of terms like *big/groß & small/klein*. Such expressions relate to *properties* rather than spatial relations or locations. Herskovits (1986) deals with the wider class of locative expressions, focussing on prepositions, and addressing dimensional (in her terms, projective) terms specifically in a separate chapter.

The various types of dimensional terms, i.e., spatial locative and directional terms and temporal terms, are schematically depicted in Figure 1 above. Here and in the following figures, German realisations are marked grey for clear distinction. Realisations focus on locative terms since these are primarily targeted here.

As Wunderlich & Herweg (1991:760) note, locatives are usually **relational** (at least "**zweistellig**" or "two digit", i.e., requiring two participants) since they are used to localise objects relative to other objects. This observation also applies to temporal dimensional terms, but not necessarily to directional ones: expressions such as *nach rechts* require only one participant from which the direction is derived (using information about the participant's position and orientation). Furthermore, some spatial terms such as *between* require more than two participants. The present work focuses on relational, predominantly locative spatial-dimensional expressions ('dimensional' for short if there is no need to be contrastive or specific). These can be further subdivided with respect to the dimensions they express. There are several reasons, which will gradually become apparent, to first distinguish between horizontal and vertical dimensions, and then subdivide the horizontal dimension into the lateral and the frontal (Figure 2):

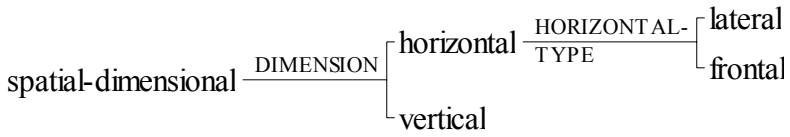


Figure 2. Spatial dimensions

1.2 Syntactic range of locative spatial-dimensional terms

In the area of dimensional terms, Herskovits (1986:156) lists a range of linguistic options, noting that most are composite expressions. She points out that their usage involves a **frame of reference** comprising basic **half-line axes** and a **reference object** (relatum). These components will be addressed separately below since they involve complex concepts interacting with real-world factors. Figure 3 below shows Herskovits' list as realisations in a schematic representation.

Herskovits' list only aims at capturing preposition groups in English. It will be taken as a starting point, with the aim of expanding it as far as possible throughout this work, including German terms and other syntactic forms. Note, furthermore, that Herskovits here includes prepositions that are neutral with respect to the half-axis chosen, e.g., *at the side of*. It is unclear why, when such expressions are included, others such as *beside* and *next to* are missing. Subsequently, Herskovits focuses on the horizontal plane. This parallels the present work, in which I start from the horizontal plane because of the obvious effects of gravity, which often preclude the localisation of one object relative to another on a vertical scale (excepting support relations). Nevertheless, I will try to account for the vertical dimension as completely as possible.

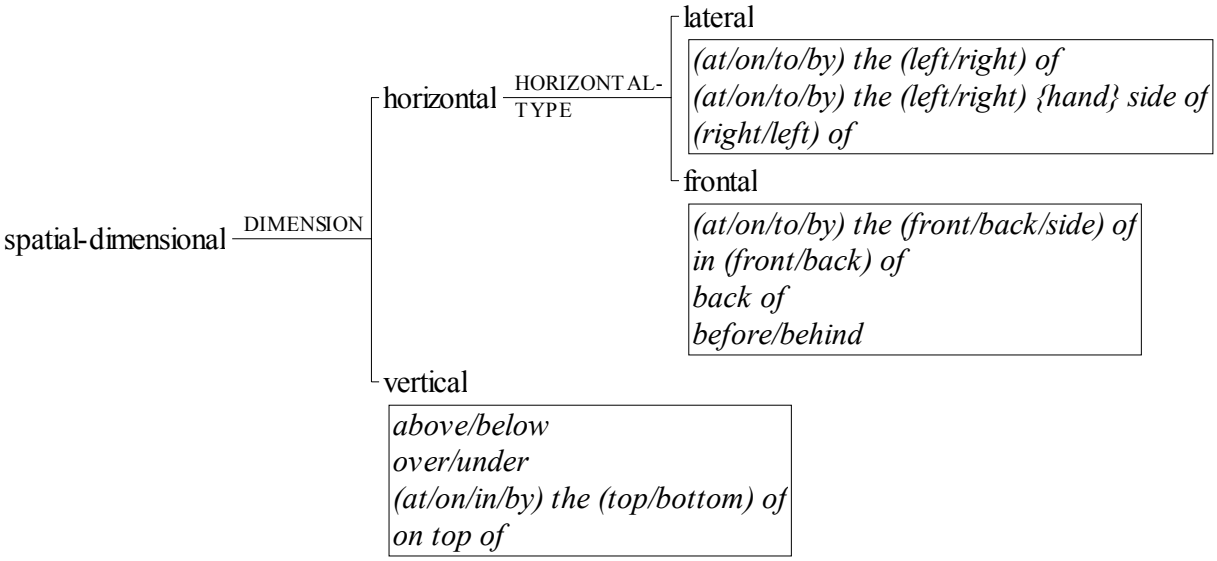


Figure 3. Spatial dimensions: Realisations in English (acc. to Herskovits)

Herskovits analyses the internal structure of the listed preposition groups as follows (1986:157):

> "Once a frame of reference is inferred, one of the base axes (or two axes, left and right, in combinations with *side* and no restricting qualifier) is selected according to which morpheme the preposition includes: *front, back, right, left, side, -fore,* or *–hind.* The remaining morphemes constituting the preposition (*at, on, in, be-,* etc.) signal different ways of constraining the location on or near the axis chosen, and some subtle but distinct contextual conditions of use."

Thus, the dimensional term denoting the axis is combined with further morphemes that systematically constrain application (see section 2.7 below).

Complementing Herskovits' analysis which is focused on preposition groups, further syntactic variability needs to be accounted for. Wunderlich & Herweg (1991), who propose a language-independent classification (based on several cross-linguistic examples), list nouns, adpositions (prepositions and postpositions, the latter of which rarely occur in English or German), adverbs and verbal particles. Locative nouns in English and German correspond to object parts as already noted by Herskovits (see above) or are used metaphorically, as in (Wunderlich & Herweg 1991:760):

(1) Die Truppen sammelten sich *im Rücken* des Gegners.
[lit: The forces concentrated at the back of the enemy.]

Locative adverbs can serve the same function as prepositional phrases, i.e., they can replace them, as in (Wunderlich & Herweg 1991:761):

(2) Anna steht *vor dem Haus*.
[Anna is standing in front of the house.]

(3) Anna steht *davor*.
[lit: Anna is standing there-in front of; "Anna is standing in front of it."]

Since locative adverbs are semantically relational but do not take a syntactic object, they are used for the implicit localisation of objects relative to contextually given relata.

Verbal particles are lexicalised parts of verbs (mostly motion verbs) that carry relational meaning, such as *aufspringen* and *hineingehen.* In some cases the motion verb itself (rather than a separate morpheme) carries relational meaning, as in *enter.*

In German, case differentiates between directional prepositional phrases (such as (4)) and locative ones (such as (5)) (Wunderlich & Herweg 1991:762):

(4) Die Katze springt auf den Tisch.
[The cat is jumping onto the table.]

(5) Die Katze springt auf dem Tisch.
[The cat is jumping on the table.]

In English, this difference is expressed by the choice of prepositions (*onto* vs. *on*). Eschenbach (2004) presents a table listing the whole syntactic range of dimensional terms (both directional and locative) which highlights several observations: First, the vertical dimension (which is not in the center of attention in the present work) offers a much wider range of terms in German than the two horizontal ones do. Second, there is a much greater range of variation in directional terms than in locational ones. Third, with respect to dimensional terms the just mentioned case distinction applies for the vertical (*über, unter*) and frontal (*vor, hinter*) axes but not for the lateral one. In the lateral axis, prepositions appear with a further case, namely, the possessive (*rechts/links des Hauses*), which is also possible with the vertical axis (*oberhalb/unterhalb des Hauses*). Adjectives and adverbs can appear with all axes. These distinctions are represented in Table 1 below (omitting Eschenbach's list of directional terms).

Note that Eschenbach's list does not include adverbial forms with additional deictic morphemes such as *davor*, nor nouns such as *(die) Front* as mentioned by Wunderlich & Herweg (cf. above), nor verbal particles (such as *vorgehen*). Eschenbach's list may well reflect the intuition that the expressions listed should be the most common ones in German, as opposed to English where the noun *front* frequently appears in a dimensional preposition group, as analysed in detail by Herskovits (cf. above). Note, for instance, that the frequently used English preposition group *to the left* can be translated unproblematically into German, yielding the rather archaic expression *zur Linken* which – interestingly – is listed by neither Wunderlich & Herweg nor by Eschenbach. Klabunde (1999:152) additionally mentions *rechterhand* and *linkerhand* (similar to *right/left hand side* in English) which are semantically similar to the adverbs *rechts/links*, but "their use is confined to a humanlike origo".

Table 1. Syntactic and dimensional variability in German locatives (acc. to Eschenbach)

Syntax	UP	DOWN	FRONT	BACK	RIGHT	LEFT
Prep [+dat]	*über*	*unter*	*vor*	*hinter*		
Prep [+poss]	*oberhalb*	*unterhalb*			*rechts*	*links*
Adjective	*ober*	*unter*	*vorder*	*hinter*	*recht*	*link*
Adverb	*oben*	*unten*	*vorne*	*hinten, zurück*	*rechts*	*links*

1.3 Core semantics for locative spatial-dimensional expressions

The aim of the present section is to work out to what extent we can identify a **core meaning** for locative spatial-dimensional expressions that abstracts from individual discourse situations as far as possible. Herskovits (1986:3) illustrates a very similar idea as follows:

> "I will show that one can make sense of many apparent irregularities of semantic data by assuming that word meanings are defined in an ideal world – in the spatial domain, a world of lines, points, surfaces, and of definite relations of inclusion, contact, intersection, and so forth. To describe and communicate facts about the complex and imperfect world that surrounds us, we must bend and stretch these ideal concepts. Obviously, this bending and stretching cannot be done at the speaker's whim; while it may not all be the result of linguistic rules, it is constrained by the need to preserve mutual comprehension, so that our references to objects and facts are, usually, successful. For that reason, ideal concepts must be complemented by a precise account of the way in which such concepts are distorted in a linguistic and situational context."

In principle, I agree with Herskovits' approach in that I assume a general distinction between semantic and pragmatic levels. However, unlike her I do not consider the semantic core meanings as 'ideal', nor do I view their variability in actual discourse as 'distortions'. Rather, I

assume that there is some kind of semantic core that pertains to every use of the terms regardless of context. This implies that the semantic core needs to be sufficiently abstract to hold in spite of the many ways in which discoursal factors can influence applicability. In other words, there is no ideal – or logically accurate – meaning that could be (in terms of Bennett 2002) "approximated" in real usage, but the term is sufficiently flexible to be used in a great variety of contexts without loss of its semantic core. Discoursal conditions, pertaining either to the discourse itself or to external factors of the situation, influence the application of spatial terms in various ways, sometimes constraining the applicability, sometimes highlighting specific aspects of the core meanings while neglecting others. Like Herskovits, I assume that such processes are regular and should be accounted for as systematically as possible. Thus, apart from this evaluative difference Herskovits' approach is a suitable starting point for present aims. In Bateman & Farrar (2004:74f.) a similar view is expressed as follows:

> "(…) in all cases the degree to which the term fits or not depends on the discourse involved, which is another way of saying that it depends on the function and purpose of attempting the description in the first place."

Thus, it is not the core meaning of the dimensional terms that is affected by diverse discourse factors, but rather, application of the term can be more or less appropriate depending on context. Another way of putting this idea is in terms of degrees of **compatibility** with the spatial situation (Freksa 1999). Freksa also points out that the linguistic alternatives that may be available in the discourse further influence the acceptability of a specific term (cf. section 2.3 below).

Discourse functions influence the applicability as well as the interpretation of terms, for example, in matters of granularity (cf. section 2.5 below) and with regard to the functionality of the entities related by the dimensional terms (section 2.8), highlighting specific aspects while treating others as irrelevant. A similar approach is taken by Coventry & Garrod (2004a) who point out (p. 33) that lexical entries may well only reflect a minimal part of what actually comes into play in usage –

> "so long as there are systematic principles to explain how the lexical entry can be flexibly bent and stretched in a context."

In approaching the nature of the ideal meaning of spatial terms, Herskovits (1986:39) starts by proposing the following:

> "The ideal meaning of a preposition is a geometrical idea, from which all uses of that preposition derive by means of various adaptations and shifts. An ideal meaning is generally a relation between two or three ideal geometric objects (e.g., points, lines, surfaces, volumes, vectors)".

Note that this view corresponds largely to that taken in the cognitive linguistics tradition where the nonspatial usages of originally spatial expressions are viewed as metaphorical derivations of basic (abstract, geometric) meanings (called proto-scenes, for instance, in Tyler & Evans 2003). A general specification of the semantic core of *dimensional* terms that is compatible with Herskovits' approach is formulated by Grabowski (1998):

> "Allgemein lässt sich zunächst sagen, dass dimensionale Präpositionen eine räumliche *Relation* zwischen einem Zielobjekt und einem Bezugsobjekt spezifizieren; die Bedeutung der einzelnen Präpositionen betrifft die Art dieser Relation."
>
> [In general, we can start by saying that dimensional prepositions specify a spatial *relation* between a goal object and a reference object; the semantics of each single preposition concerns the nature of this relation.]

Furthermore, he points out that the arguments of this relation can either be concrete objects or the spatial locations (places) that these objects occupy (Klein 1991), which seems to be a matter of theoretical stance and modelling decision.

Herskovits (1986:181) lists the following expressions as having "the most general conditions of application": *to the right of, to the left of, in front of, behind.* She describes the ideal meaning of these terms as "graded concepts with the denoted axis as their focal region" (p 184). The other terms used for external relationships (given above) are also based on this ideal meaning but are more constrained in application by pragmatic factors (Herskovits 1986:188ff.). Herskovits motivates the gradedness of the concepts by pragmatic arguments. In other words, the application of the terms in discourse indicates in several ways that the underlying semantics of the dimensional terms must be based on graded concepts, with the axis itself as ideal interpretation, although pragmatic factors (especially contrast, see section 5.1 below) may affect the influence of the graded structure.

This view of including the gradedness of interpretation in the semantics of dimensional terms mirrors a wealth of literature dealing with this phenomenon in the (nearly) two decades since Herskovits' work. The concept of **spatial templates**, sometimes expressed in terms of **prototypicality effects** (cf. Rosch 1978, Lakoff 1987), has been developed in detail in psycholinguistic research (see section 2.3), building the basis for developing computational models (see section 3). The experiments carried out to determine the structure of the templates are designed to trigger utterances that describe the spatial relation between a reference and a target object for a specific purpose, e.g., to find a hidden object (e.g., Zimmer et al. 1998). Thus, they inevitably interact with discourse-related factors involved in application, and they inevitably involve concrete objects whose properties, as is well known, influence application as well. For example, as Herskovits (1986:184ff.) shows, the interpretation of the ideal meaning depends on the spatial extension of the reference and target objects, and other discourse-external factors as described in section 2. These facts illustrate the difficulty of separating abstract meaning components from actual usage in discourse – there seems to be no usage-independent evidence for a basic semantics of dimensional terms. Accordingly, psycholinguistic experiments are usually not designed for isolating abstract meaning features of spatial terms but rather aim at identifying cognitive or pragmatic criteria for their usage. Often, however, the term 'semantics' is used in a number of different ways for characterising findings on the interpretation of spatial terms in different application situations.

In the present approach, it is assumed that the nature of real-world objects and pragmatic factors cannot affect the (abstract) semantics of a linguistic expression that does not directly denote these real world objects, and that is not restricted in usage to a limited range of objects. Thus, the semantic core should be precisely that part of the description of a term that remains largely unaffected by pragmatic factors as well as by the semantics contributed by the discourse context (relating, for example, to the objects involved). It is then the *applicability* of a term that is affected by such factors, not the term's semantics, which should be defined in such a way as to allow for flexible usages in various discourse contexts. In other words, the semantics of a term (at any given time, for any given language user or community, as far as this can be determined) should be restricted to aspects that can be specified without knowledge about the situation context. For the present concerns, this means that, as far as possible, factors pertaining to the *relation* between two objects (as denoted by the dimensional term) should be distinguished from factors pertaining to the *objects* that are related by the dimensional term. The former concern the semantics of dimensional terms directly, while the latter pertain to the semantics and conceptualisations of the objects concerned. In constructions such as *behind the church*, both factors jointly contribute to the interpretation, which may lead to the inference of a spatial relation that is in many ways different from the spatial relation expressed by a superficially similar construction such as *in front of the oven.*

Thus, together with further factors such as conceptualisations of configurations as expressed in terms of reference systems, availability of competing expressions (Freksa 1999), etc., object features contribute to the applicability and interpretation of the dimensional terms in

context, without however influencing their abstract semantics. This observation does not rule out the possibility that functional features pertain directly to the relation between objects, such as *support, containment,* and the like, as discussed below. These are then included in the abstract semantics of the spatial term. The abstract semantics of the objects involved, on the other hand, which may also include functional features (which may or may not match the functional features of the spatial term), affect the applicability of the spatial term. Such effects are therefore discussed (in the present work) along with other pragmatic and context-related factors.

Clear borderlines are certainly difficult to draw (if feasible at all) since spatial relationships always depend on concrete real-world objects, even though attempts are made to abstract from them, as is the case with Herskovits' 'ideal meaning' that abstracts as far as possible from specific usages. Grabowski (1998:11) notes that, in the case of dimensional terms, the definition of an abstract semantics is especially difficult

> "weil der charakterisierende, sprachinterne Bedeutungsaspekt (i.S.v. Freges 'Sinn') nur wenig Spezifikation zu erlauben scheint. Die – semantisch interessantere – große Variabilität der Verwendungsfälle, die sich in einer Mehrzahl alternativer Wahrheitsbedingungen abbildet, scheint demgegenüber nur im Rekurs auf die Situationskonzeptualisierung von Sprachverwendern erschließbar zu sein. (…) Somit stellte sich die in den verschiedenen Ansätzen unterschiedlich gelöste Frage, wo die lexikalische Bedeutung dieser Präpositionen aufhören und wo ihre (kognitiv-konzeptuelle) Kontextabhängigkeit beginnen soll."
>
> [because the characterising, language-internal aspect of the semantics (cf. Frege's 'sense') allows only for very little specification. The – semantically more interesting – greater variability of application cases, which is represented in a plurality of alternative truth conditions, in contrast, seems to be accessible only by considering the situation conceptualisation of language users. (…) Thus, a question arises which is solved differently in the various approaches, namely, where the lexical meaning of these prepositions is supposed to end and where their (cognitive-conceptual) context dependency is supposed to begin.]

In light of such difficulties, the present approach does not attempt to solve the problem of how or whether the different kinds of application are represented in the lexicon. According to Herskovits, various **use types** are part of the lexicon of spatial terms ("attached to each ideal meaning of each preposition", 1986:87), an approach that is criticised by Coventry & Garrod (2004a:28) as follows:

> "Herskovits ends up listing the exceptions in the lexicon in much the same way as is required in the minimally specified accounts we have discussed above. The difference is that she maintains that the use types are legitimised on the grounds that they relate in motivated ways to the ideal meaning. However, it is unclear what form use type representations are supposed to take and, additionally, what the role of object knowledge is in her account. Moreover, the use types are lexicalised for prepositions, and this requires that the correct use type be selected from the lexicon in each context, although how this is done is not considered."

An example for how an analysis of spatial terms completely independent of application could look is presented by Manning et al. (2002), who examine judgments of similarity between frequent English spatial prepositions presented with or without a sentence context. They identified four different dimensions of similarity, three of which were kept more or less constant across contexts. The four relevant dimensions were interpreted to reflect degrees of verticality, containment, distance, and perhaps (but less clearly) contact. However, since the investigation did not include most of the dimensional expressions (excluding *in front of, left,* and *right* because they do not appear as simple prepositions), the study is not particularly informative for present purposes.

I will not pursue the semantics/pragmatics discussion further, but simply note here that the idea of a core semantics may be useful to some degree for explanation, but it clearly has limitations and controversial associations. The distinctions drawn here might not be applicable or useful for different purposes than those pursued in the present work.

Specifically, I will view gradedness effects as important for the application of dimensional terms (see section 2.3 below), while not classifying them as part of their core semantics. This view is supported by the observation (Coventry & Garrod 2004a:29ff.) that prototype effects occur with all kinds of linguistic expressions as well as with categories that are created on the spot (i.e., which do not possess a lexical entry at all), such as *things to take from one's home during fire*. Coventry & Garrod (2004a:32) conclude that

> "the existence of prototype effects does not mean that the structure of representation in the lexicon is in terms of one (or more) prototypes."

I will furthermore not follow Herskovits (1986:190f.) in ascribing two distinct basic meanings to the dimensional terms depending on the underlying reference system, but assume that the choice of reference system is basically independent of the dimensional term itself. This assumption is motivated, first, by the fact that most dimensional terms can be used flexibly with different reference systems (except for the absolute which uses different linguistic material, see section 2.1 below), and second, by the fact that reference systems do not depend on language but can be addressed in non-linguistic experiments (e.g., Levinson 2003). Thus, in my view, dimensional terms and underlying reference systems interact but should not be conflated when describing the semantic core. This view is reflected in the lexical entries for dimensional prepositions given by Eschenbach (2004), defining the semantic core of each term relative to a reference system that is viewed as a contextual component. For instance, the semantic form of German *vor* (in front of) is described as follows (G refers to the ground object, i.e., the referent):

> "the figure is spatially included in a region that is derived from G and the principal direction FRONT as given by a reference system *RSYS*."[1]

The semantic forms of *hinter, über, unter, rechts,* and *links* can be described similarly by replacing the principal directions (half-axes) accordingly. Note that this description does not include gradedness in applicability. Eschenbach takes it as a starting point for further analysis, emphasizing that refinement is necessary in several respects: this general model does not account for the influence of functional aspects which can lead to differences between the directions, it does not explain how the directions are derived from the ground object, and it does not include the specific or contrasting requirements for other, related lexemes such as *oberhalb, unterhalb* with regard to gradedness (see section 2.3 below).

An alternative view that is equally independent of reference frames but assumes different geometric features in the semantics of terms describing relations on the lateral vs. frontal axes is proposed in Eschenbach (1999)[2]. Here, she states that "[t]he semantic core of German *vor* and *hinter* can be specified based on an oriented linear frame of reference" (p. 340), while *rechts* and *links* are based on half-planes rather than linear ordering (p. 343). Accordingly, different axiomatic formalisations are suggested for the two axes[3]. The main motivation for this distinction is the fact that *vor/hinter* can be applied for linear ordering in the case of motion, while no corresponding linear ordering for *rechts/links* is induced in dynamic situations. However, as will be seen in section 2.2.1.6, the reference frames induced by motion are specific cases of application; furthermore, motion does produce lateral as well as frontal axes: the direction of motion is always forward, i.e., defines the front axis, but resulting from that ascription of a front side the other axes can also be determined. Therefore, it is assumed here that the distinction is not as fundamental as suggested by Eschenbach. Rather, it is likely that, with respect to both axes, both linear and planar interpretations can be applied depending on context, which suggests that they do not belong in the core semantics at

[1] For a formal version of this lexical entry and those for other lexemes see Eschenbach (2004).
[2] The specifications for the different directions presented in later sections in Eschenbach (2004) are compatible with this view.
[3] Similar descriptions of the corresponding English terms are given in Eschenbach & Kulik (1997).

all. Ordering phenomena occur as a consequence of the application situation which, for instance, involves order in motion or in a queue; furthermore, the asymmetry of the frontal axis as opposed to the lateral axis is important (see section 2.1 below). Thus, while it is certainly true that the orientation of the frontal axis opens up possibilities for usage that are not available on the lateral axis because of its lack of directedness, there seems to be no empirical foundation to the claim that the semantics and applicability of the terms are fundamentally different, as Eschenbach claims.

The proposals presented so far mainly treat the basic dimensional terms representing the six spatial directions (*left, right, front, back, above, below*) similarly, focusing on the geometric aspects of the abstract semantics of the dimensional terms. In a different approach developed in detail, for example, by Aurnague & Vieu (1993) with regard to French terms, the geometric level represents only part of the semantic description. Another part is the contribution of a functional element, such as SUPPORT in the case of the topological term *on*, or CONTAINMENT in the case of *in* (Garrod et al. 1999). According to Aurnague & Vieu, the functional element also belongs to the abstract semantics because it is valid across discourse contexts. Pragmatic aspects pertaining to individual (kinds of) discourse situations are then analysed on a third level. Coventry & Garrod (2004a) propose a systematic account of this view, pointing out how the two elements interact in the application of spatial terms. For example, the lexical entry of *in* contains a functional component of location control and a geometric component of enclosure, which come into play to different degrees in different discourse contexts (2004:53):

> "When objects exhibit only weak enclosure (…), then firm evidence of location control is required to support an *in* relation. (…) However, when the evidence for location control is weak, there needs to be a strong geometric enclosure to support the *in* relation. The same is true for the *on* relation."

In the present work, a distinction is made between associated functional features of the spatial expression itself (or the spatial relation that is represented by the term), and functional aspects that pertain to the objects involved. Thus, an expression like *in* is associated with the functional element of containment because it is restricted to spatial relationships of containment. Therefore, the expression can only be used with objects that allow for such a relationship. This inherent feature of *in* explains the difference in acceptability between sentences using *in* together with different objects:

(6) The cup is ?in / on the table.

(7) The hole is in the table.

(8) The cup is in the bowl.

Since objects can be associated with different kinds of functional features, naturally different objects can match with the spatial term to different degrees, also depending on the situational context. These aspects constitute the major part of research concerning functional geometry (a term proposed by Garrod et al. 1999) since it is usually the function of the objects that contributes the evidence for location control and the like. This field of research is addressed separately in section 2.8 below. It will be seen there that the functions of objects can influence the applicability of spatial terms in a number of ways, going beyond the (usually fairly clearcut) functional element of the spatial term itself.

Here, I will take a closer look at the work of Tyler and Evans (2003) who propose a similar view as that taken in Aurnague & Vieu (1993), but deal exclusively with English expressions. Their analysis draws on, reviews, and integrates previous accounts, including the famous ground-breaking work on *over* by Brugman (1988). In their terms, the semantics of each spatial term is described by a spatial "proto-scene" together with a functional element that serve as basis for all additional senses derived from this basic meaning. Such a proto-scene abstracts away from the details of specific spatial scenes but consists of a "schematic trajector" (TR) and a "schematic background element", a landmark (LM). Thus, their

description of the vertical dimensional term *over* is that "the TR is higher than but within potential contact of the LM" (p. 65). The functional aspect involved here is that either one of the participants "is conceptualized as being within the sphere of influence" (p. 66) of the other. This description is suited to differentiating between *over* and the very similar term *above* that nevertheless has different areas of application. Tyler and Evans (2003:66) claim that this is because "the proto-scene for *above* designates a spatial configuration in which the TR is higher than but not in potential reach of the LM". Furthermore, they propose a methodology which they use to demonstrate which kinds of senses have evolved as distinct from the earlier semantics of a term, and which interpretations arise from the context in which the terms occur. For present purposes, derived senses are less relevant because spatial settings are in focus here, as opposed to Tyler and Evans' work where specifically non-spatial interpretations of originally spatial terms are highlighted. But not all derived senses are non-spatial, as illustrated by the following sentence, which exemplifies the "On-the-other-side-of Sense" (Tyler and Evans 2003:81):

(9) Arlington is over the Potomac River from Georgetown.

This sense, according to Tyler and Evans, evolved from a reanalysis of a frequent interpretation of the original proto-scene in which the trajector is in motion, first upward and above the landmark, and then – because of the forces of gravity – downward at the other side of the landmark, as in (Tyler & Evans 2003:71):

(10) The cat jumped over the wall.

The authors propose that this "above-across" interpretation should not be analysed as a distinct sense of *over* because of the contributions of the verb plus world knowledge (a cat cannot hover above a wall for an extended length of time). Such contributions of other participants of a sentence also lead to the distinct interpretations involved in (10) versus (11) below (Tyler & Evans 2003:72): here the latter case involves a transversal interpretation that is contributed by the bridge, not by *over*, which itself (according to the authors) does not contain a path component in its central sense. This is because bridges "typically facilitate passage across obstacles such as rivers" (p.73).

(11) The girl walked over the bridge.

The authors then claim that the "On-the-other-side of Sense" as in example (9) is derived from these non-distinct senses of *over* and has become distinct, since sentences such as (9) do not necessarily involve motion (though they may still evoke associations of motion).

Nevertheless, this additional spatial sense highlighted by Tyler and Evans' work is not directly relevant to present purposes since it no longer directly involves the vertical dimension present in the original sense of *over*. Thus, while Tyler and Evans' work (which builds on a long tradition of work in cognitive linguistics, e.g. Langacker 1987/1991, Talmy 2000) highlights the fact that dimensional terms can be employed in non-dimensional as well as completely non-spatial contexts, from now on I will concentrate on the purely dimensional senses, which are usually directly reflected in Tyler and Evans' analyses of proto-scenes for each (dimensional) term. Here, the most interesting aspect is the functional interpretation that goes with each proto-scene. Table 2 shows an overview[4] of the basic senses ascribed to dimensional terms by Tyler and Evans (2003)[5]. The authors do not consider the lateral axis (*left, right*) at all, presumably because non-spatial senses (which are the main focus of their work) can scarcely be derived from them.

[4] Unfortunately, Tyler and Evans do not provide any comprehensive overview of this kind; rather, the relevant information is distributed throughout the book. I hope that this overview does not misinterpret their proposal.

[5] I do not agree with Tyler and Evans' proposal that (synchronically) the semantics of *before* and *after* is basically spatial. However, since it is uncontroversial that these terms can in fact be used spatially the interpretations are included here.

The authors point out that, with regard to the basic senses, pairs of spatial terms often constitute a **contrast pair**, as is the case with *above/below*, *over/under*, *up/down*, etc. Thus, the spatial description involves a contrast with regard to the direction on the same axis, while the functional element in each case is either equivalent (as with *above/below*) or contrastive (as with *up/down*).

Table 2. Basic senses of dimensional terms (acc. to Tyler and Evans)

Expression	**Spatial Proto-Scene**	**Functional Element**
over	the TR is higher than but proximal to the LM	potential reach ("sphere of influence") between TR and LM
above	the TR is higher than but distal to the LM	no potential reach between TR and LM
under	the TR is lower than but proximal to the LM	potential reach ("sphere of influence") between TR and LM
below	the TR is lower than but distal to the LM	no potential reach between TR and LM
up	an oriented TR is directed towards the top of an oriented LM (vertical axis)	positive value: physically elevated means more visible, more accessible, in a normative position, etc.
down	an oriented TR is directed towards the bottom of an oriented LM (vertical axis)	negative value
in front of	an oriented LM is directed towards an unoriented TR (frontal axis)	(perceptual) accessibility
before	an oriented TR is directed away from an oriented LM ("in tandem" alignment) (frontal axis)	leading, sequence
behind	an oriented LM is directed away from an unoriented TR (frontal axis)	lack of perceptual access
after	an oriented TR is located at the back of an oriented LM ("in tandem" alignment) (frontal axis)	following, (intentional) pursuit

For many of the terms, it is proposed that the semantics involve the orientation of either the landmark or the trajector. However, especially in the case of *in front of/behind* this interpretation runs into problems. Tyler and Evans propose that cases like example (12) are to be viewed as exceptions (p. 161):

(12) Sarah stood in front of the bush.

Here, it is the orientation of the trajector rather than the landmark that is responsible for the conceptualisation of a frontal region. While the authors' explanation of this phenomenon may be controversial, a decisive problem with their account lies in the fact that, altogether, they only account for examples involving intrinsic reference systems, where either the trajector's or the landmark's orientation is used for reference. Only in a footnote on page 162 do they mention the possibility of the vantage point of a third entity. In the present approach I rely on Levinson's classification of reference systems (see section 2.1 below) which neutrally allows for different vantage points along with a variety of possible reference systems, neither of which are inherent parts of the semantics of dimensional terms. This classification will show that it is only in the case of intrinsic reference systems that the orientation of the entities

involved matters. In relative reference systems, the observer's vantage point takes over this function. Furthermore, Tyler and Evans' account clearly underestimates the role of orientation of both trajector and landmark in the application of all dimensional terms: while it is clear that the landmark's intrinsic front is decisive for assigning spatial axes in the application of intrinsic reference systems (see section 2.1 below), recent research (Burigo & Coventry 2005) indicates that even the trajector's properties and orientation affect application of vertical terms. Nevertheless, there is no reason to assume that such effects belong in the abstract semantics of dimensional terms.

The association of accessibility with the frontal axis is mirrored in various other accounts; Eschenbach (2004), for instance, assumes that the front region of a person (in an intrinsic reference system)

> "is the region of his or her best access by vision, manipulation, and motion. The region behind the person is the opposite region, which can be accessed less well. (…) Relative uses of *vor* and *hinter*, in contrast, deal with distance and occlusion."

Thus, in her account the semantics of the expressions of the frontal axes are associated with the different effects of the reference systems on accessibility.

Grabowski (1998) even proposes that all horizontal expressions are related to perceptual accessibility from the origin's position. His definition of the abstract semantics of the four horizontal terms goes as follows.

> "*vor:* Es gibt eine anthropomorphe Origo. Das Zielobjekt befindet sich in der Region des Bezugsobjekts; es befindet sich auf der von der Origo aufgespannten Ersten Horizontalen, und es ist für die Origo perzeptuell zugänglich. (…)
>
> *rechts:* Es gibt eine anthropomorphe Origo. Das Zielobjekt befindet sich in der Region des Bezugsobjekts; seine Position auf der von der Origo aufgespannten Ersten Horizontalen entspricht ungefähr der Position des Bezugsobjekts auf der Ersten Horizontalen. Seine Position auf der Zweiten Horizontalen weicht von der Position des Bezugsobjekts auf der Zweiten Horizontalen in Richtung der rechten Seite der Origo ab. – Mit anderen Worten: Wenn die anthropomorphe Origo nach rechts 'blickt', kann sie das Zielobjekt sehen."
>
> [*vor (in front of):* There is an anthropomorphic origin. The goal object is located in the region of the reference object; it is located on the frontal axis starting from the origin, and is perceptually accessible for the origin. (…)
>
> *rechts (right):* There is an anthropomorphic origin. The goal object is located in the region of the reference object; its position on the frontal axis starting from the origin corresponds approximately to the position of the reference object on the frontal axis. Its position on the lateral axis departs from the position of the reference object on the lateral axis into the direction of the right side of the origin. – In other words: If the anthropomorphic origin 'looks' to the right, it can see the goal object.]

The definition of *hinter* is equivalent to that of *vor*, except that the trajector ("Zielobjekt") is perceptually not or only partly accessible for the origo. The description of *links* mirrors that of *rechts*. Thus, in Grabowski's view the spatial relation between the two objects is less relevant than perceptual factors:

> "Insgesamt lassen sich die dimensionalen Relationen (der Ebene) und ihr sprachlicher Ausdruck durch dimensionale Präpositionen somit einheitlich als Fälle der perzeptuellen Zugänglichkeit eines Zielobjekts für eine antropomorphe Origo bestimmen. Unter dem Ziel einer einheitlichen semantischen Beschreibung erweisen sich somit (überraschenderweise) die *Perzeptions*bedingungen der Origo als bedeutungsrelevant, nicht hingegen die relative *Raum*lage des Zielobjekts zum Bezugsobjekt."
>
> [Thus, altogether we can define the (planar) dimensional relations and their linguistic expression by dimensional prepositions uniformly as cases of perceptual accessibility of a goal object for an anthropomorphic origin. With the aim of a uniform semantic description, the *perceptual* conditions of the origin thus turn out (surprisingly) to be semantically relevant, rather than the relative *spatial* position of the goal object to the reference object.]

Specifically with regard to *hinter,* where a lack of perceptual accessibility is proposed as part of the semantic core, this description is compatible with Tyler and Evans' account. It remains to be seen whether perceptual accessibility is really as fundamental to the core semantics of the dimensional terms as proposed, or whether in actual discourse situations occur in which it does not play any role. If that is the case, the perceived importance of perceptual accessibility should be viewed as conceptually derived from a typical situation in which the spatial arrangement in fact supports or impedes perceptual access, rather than as an inherent part of the terms' semantics.

Features such as perceptual accessibility, association (of the vertical axis) with gravity, and (sometimes) hand preferences for the right half axis (in contrast to the often somewhat weaker left hand), are described as functional elements of spatial terms, e.g., in Aurnague & Vieu (1993), Tyler & Evans (2003), and Eschenbach (2004). Common to those approaches is the requirement that the associations should hold across discourse situations. This should not be confused with functional features of the objects involved, as described in section 2.8 below.

In the following, I will address relevant factors influencing the application of locative dimensional terms separately, keeping in mind that it is often enough a matter of theoretical decision to include some of these aspects in the lexical semantics of the terms, unless it is clearly the objects, not the relation between them, that influence the terms' application.

2 Discourse-external factors involved in the application of spatial dimensional terms

2.1 Axial structure in language and cognition

Unlike time which involves only one dimension and therefore only one basic axis, spatial concepts are applicable in three different dimensions, complicating representation. Typical semantic descriptions of dimensional terms, such as the proposal by Eschenbach (2004, see p 10), involve the notion of **axis** (or **direction** in a different terminology). More concretely, interpretation of the terms builds on the conceptualisation of regions surrounding three basic axes, corresponding to each of the three dimensions. Various ways of representing spatial information in language and other modes (by way of axes, half-axes, half-lines, vectors, regions, etc.) capture basically this same idea, as shown, for instance, in the diverse contributions in van der Zee & Slack (2003).

In principle, it is also possible to conceptualize the spatial directions as half-planes rather than axes, dividing three-dimensional space into half-spaces, rather than two-dimensional space into half-regions. In language use and in most literature on the topic, however, this conceptual option does not seem to play a major role. Even if the concept may be underlying some linguistic representations, this is usually not reflected in the language, as is exemplified with respect to various interpretations of one utterance in section 2.5 below. In other words, most linguistic representations that employ the dimensional terms can be interpreted as though referring to a two-dimensional rather than a three-dimensional world, in which case it is sufficient to talk about axes separating half-planes, rather than planes separating half-spaces. This intuition is supported by the observation that usually a maximum of two, not three, dimensional terms are combined in one single utterance. This implies that, usually, the third dimension does not play a role in the spatial relationship represented by the utterance. According to Herskovits (1986:76) spatial relationships are often described ignoring the vertical plane:

> "Horizontal coordinates stand out in our experience of objects, because we are normally bound to the ground, and move on it to approach objects, so we often ignore altitude".

Most empirical work on dimensional terms, furthermore, is restricted to two dimensions by the spatial setting: usually, the vertical dimension is missing, except if pictures are involved, in which case the frontal axis is neglected.

With regard to the cognitive status of the three axes, some differences can be noted. For instance, the vertical axis is viewed as special because it is defined on the basis of gravity, and thus independently of the orientation of the other axes (e.g., Nikanne 2003). Psychological and psycholinguistic research (e.g., Bryant et al. 2000, Landau 2003) also shows that the vertical axis is cognitively privileged, as proved by results in memory and attention tasks, language acquisition, etc. Furthermore, the frontal axis has precedence over the lateral one because it is directed and derivable from objects' fronts, while the lateral axis is symmetric and needs to be derived from the orientation of the frontal axis. Retz-Schmidt summarises the results and arguments of earlier work as follows (1988:97):

> "[T]he vertical dimension has the status of a privileged direction because it is fixed by the gravitation of the earth, whereas in both horizontal dimensions, man [sic] can move freely. Thus, the vertical dimension can be conceived of as the primary dimension. Moreover, the human body is asymmetric in the front-back dimension, enabling man to distinguish more easily between front and back than between left and right. This asymmetry permits the less salient distinction between the front-back dimension as the secondary dimension and the left-right dimension as the tertiary dimension".

Ehrich (1985:132) notes that this symmetry of the lateral axis is reflected in language:

> "Die laterale Symmetrie des äußeren Körperaufbaus hat auch einen Reflex im Aufbau der sprachlichen Raumbegriffe: *neben x* bedeutet 'an einer der beiden Seiten von x', und es bleibt offen, ob es sich um die linke oder die rechte Seite handelt, m.a.W. *neben* spezifiziert eine Dimension und lässt die Orientierung offen. Die beiden anderen Dimensionen lassen keine derartige Neutralisierung zu, es gibt keine Entsprechung zu *neben* (im Deutschen ebensowenig wie in anderen Sprachen), die zwar Vertikalität oder Horizontalität spezifiziert, aber nicht explizit macht, ob etwas vor oder hinter, über oder unter x plaziert ist."
>
> [The lateral symmetry of the outer body structure also has a reflection in the structure of the linguistic spatial expressions: *neben x (beside x)* means 'at one of the two sides of x', and it remains open whether this concerns the left or the right side; in other words, *neben* specifies a dimension without specifying the orientation. The two other dimensions do not allow for such a neutralisation; there is no counterpart to *neben* (neither in German nor in other languages) which specificies verticality or horizontality without at the same time making explicit whether something is located in front of or behind, above or below x.]

This observation provides a good reason why Herskovits (1986) and several other authors intuitively (and often without further explanation) include *neben, beside* and sometimes further terms in the description of dimensional (projective) terms, which is not self-evident since the terms depart in several respects (e.g., morphologically) from the other expressions.[6]

Franklin et al. (1995) specifically examine the relationship between the frontal and the lateral axes in surrounding space (i.e., the horizontal plane in relation to the human body). Although Retz-Schmidt's summary above shows how the salience of the human body has long been recognized in the literature, previous psychological work addressing spatial regions often dealt with abstract pictures rather than surrounding space. The results of such work then showed no difference between the two horizontal axes. In contrast, the work of Franklin et al. (1995) proves the cognitive precedence of the frontal axis in surrounding space in several respects. For example, frontal regions were treated as larger than others (see section 2.3 below), resolution (i.e., discriminability of directions) inside the regions was higher, and linguistic reference centered around the frontal axis.

[6] Note that Latin has a term used for the vertical dimension in a direction-neutral manner, namely, *altus* (high, deep). This term however differs from those considered here in that it is attributive rather than relational.

Landau & Jackendoff (1993:230) specify the relationship between axes and language as follows:

> "The three principal axes can be viewed as extending from the center of the reference object to provide six possible directions. Centered around each half-axis is a region that defines the acceptable space for different prepositions."

Thus, we can either talk about three *axes* (lateral, frontal, vertical) or six *half-axes* or *directions* (left, right, front, back, up, down). Herskovits (1986:199) opts for using "half-line axes" for the following reason:

> "A half-line axis is a straight line starting at some point and extending to infinity in one direction only. A description of the frame of reference in terms of six half-line axes is more convenient than one in terms of three oriented axes extending to infinity in both directions, since the composite prepositions make direct reference to six directions, although opposite axes are not always symmetric".

While pointing to the importance of axial structure in both language and cognition, Landau (2003) finds that a specification of half-axes, i.e., direction information, has more effects in linguistic tasks than in memory. Thus, people distinguish positive and negative ends of vertical and horizontal axes in language but not necessarily in non-linguistic tasks. She concludes (Landau 2003:36):

> "The facts are consistent in suggesting that direction is represented separately from the axes themselves, and the evidence from spatial impairment suggests that direction may be a more fragile component of spatial representation than axial structure."

Since the relationship between axes and half-axes is hyponymical it is convenient (for present purposes) to refer to the basic axes whenever the directionality is irrelevant, and to the half-axes (or directions) and their linguistic labeling whenever a difference becomes apparent, such as in the actual usage of a single term.

The **acceptable space** (the region of applicability) of dimensional terms is based on the relationship of the referent to the relatum with respect to one of the basic axes. The **gradedness** of acceptability then concerns the degree of deviation from these axes (section 2.3). The frame of reference (section 2.2) determines whether the relatum's inherent axes are used in the application of the dimensional term (in the case of an intrinsic reference frame) or an axial structure is imposed externally (in the case of a relative reference frame). In both cases, the notion of axis is usually understood (implicitly) as an abstract linear structure without specifications of extensions in other dimensions. Some observations concerning the treatment of extended objects is discussed specifically in section 2.4 below.

The former case (using the relatum's inherent axes) raises the question of how an object's internal axes are derived. In natural interaction, these may as such be less relevant than their (top, bottom, front, back, left, right) **parts** or regions, which systematically interrelate with the underlying linear axes. As Landau & Jackendoff (1993:220) point out, such object parts are not subsidiary in the way handles or wings are; instead, "they denote regions of the object based on its inherent orientation".

In order to determine an object's internal regions, several different aspects come into play, most of which are not physically determined but rather based on human cognition and behavioural conventions (Herrmann 1990:120). First, in animate beings the part containing perceptual organs defines the intrinsic front. Herskovits specifies the case of a human being's axial structure in a prototypical situation (1986:157):

> "A human being in canonical position defines six half-line axes with origin at him/herself. The down axis follows the direction of the gravitational field and the up axis is opposed to it. The four other axes are in an horizontal plane: the direction in which the observer's eyes are facing defines the front axis, and the back axis is opposed to the front one; right and left axes are orthogonal to the front-back line."

Note that this definition is rather vague: it is not clear whether the origin should be understood as an abstract point (in which case the exact position of this point would need to be clarified), or a structure resembling the human's body; the view direction is also not very clearly defined. However, for analysing natural language such a vague definition is sufficient because language is generally not used in a mathematically precise way. The situation context usually provides further information that enables successful communication in spite of such underdeterminacy.

In other cases, typical direction of motion and orientation to the observer play a role (Miller & Johnson-Laird 1976:400f.). Thus, arrows have an intrinsic front because of their associated direction of motion, and a desk's front is defined by its usual orientation to a person. Intrinsic tops are dependent on an entity's characteristic orientation on the vertical axis. Levelt (1996:87) points out that the attribution of intrinsic sides is not always uncontroversial, and can lead to problems in communication:

> "The felicity of speaker/hearer coordination in the intrinsic system is (…) crucially dependent on the shared image of the relatum. First, coordination in the intrinsic system is only possible if the relatum is oriented. (…) Second, frontness is an interpretative category, not a strictly visual one."

Often, functional features come into play that are derived from our characteristic uses of the objects. These are also decisive for the determination of reference regions (see section 2.8 below).

With natural entities, it has often been pointed out that axes are not equal (e.g., Marr 1982). Landau & Jackendoff (1993:221) distinguish between one **generating** axis (the object's principal axis, e.g., the vertical axis in humans) and two **orienting** axes orthogonal to each other and to the generating axis (the lateral and the frontal ones). These axes can be marked as **directed** or **symmetric**; the vertical and frontal axes in humans are directed, while the lateral one is symmetric. In contrast to this, consider an arrow which "has a directed generating axis but no significant orienting axes" (Landau & Jackendoff 1993:221). These distinctions are responsible for the ascription of object parts, since only directed axes can have one preferred direction of typical encounter or motion, which define object fronts.

Due to the inherent orientation of the frontal axis[7], it is possible to conceptualise a linear order with respect to entities positioned on that axis. Thus, people situated closer to the front end of a queue, for example, are viewed as prior in that order. However, ordering situations apply primarily when time is involved. In the queue example, people waiting closer to the front end in a queue will be served *earlier than* people behind them. Thus, sequentiality here corresponds to an order in time as well as in space. Further situations in which spatial scenes can be conceptualised as linear are motion events, and linguistic representations of complex spatial situations, e.g., in localisation sequences and route descriptions. All these phenomena in some sense involve time as well as space and are not considered here in detail (but see Tenbrink, in prep).

Once the intrinsic front has been determined, the back side is straightforwardly inferred (by its position at the opposite side), but not necessarily the lateral directions. Here, choice of perspective is decisive. Objects can be viewed from the outside, such as a picture, where the *left side* is defined by the observer's position. But some objects can also be viewed from the inside, as is usually the case when talking about the internal parts of a person, but it can also be true for objects that can be occupied by a person, such as a car. This difference is often referred to as **handedness** in the literature (e.g., Pribbenow 1992:25, Levinson 2003), a term which relates to features of the objects themselves rather than to the perspective taken by an

[7] Although the vertical axis is also viewed as oriented, ordering information for some reason seems almost exclusively to pertain to the frontal axis, though comparisons can also occur in the vertical dimension (using *higher, lower* and the like).

observer. Likewise, Herrmann & Grabowski (1994:110) distinguish **Vehikelobjekte** vs. **Gegenüberobjekte**. Among the former are chairs, cars, clothing, etc., among the latter mirrors, desks, cupboards, and buildings. Retz-Schmidt (1988:98) points out how the objects' internal sides and regions are determined systematically, namely, by viewing them either from the inside (if feasible) or from the outside:

> "[I]f the reference object is seen from the outside, the left side is clockwise from the front (…), whereas if it is seen from the inside, the left side is counterclockwise from the front (…). Human beings and other animate beings (with a perceptual apparatus that defines their fronts) as reference objects are also treated as if they were seen from the inside".

The two object categories are thus distinguished by the way they are *typically* viewed, which seems to be uncontroversial.

A different account of the two possible ways of axis ascription is proposed by Grabowski (1999:138). He postulates that not the lateral axis is ascribed in different ways, but the frontal one. The lateral sides are in all cases those of the observer:

> "[D]ie Dimensionierung und Polarisierung der Zweiten Horizontalen erfolgt in allen Fällen einheitlich durch die anthropozentrische Zweite Horizontale der jeweiligen Origo-Instanz."
>
> [The dimensionalisation and polarisation of the lateral axis takes place uniformly in all cases by the anthropocentric lateral axis of the particular origin.]

In this view, the interesting question is why the front side of the object (which is responsible for the usage of *front*) is sometimes the one close to the observer, and sometimes the more remote one. Grabowski's proposal is that this can be solved via the notion of perceptual accessibility (cf. Grabowski 1998, see p. 14). He describes the relation as follows (Grabowski 1999:148):

> "Die Objektseite, die in Richtung der Sichtlinie der Origo weist (Innenperspektive) oder die Objektseite, die von der Origo aus an dem Objekt sichtbar ist (Außenperspektive), erhält das Attribut 'Vorderseite'. Die Frontzuschreibung ist somit ein aus den Perzeptionsbedingungen der Origo *abgeleitetes* Phänomen. Wenn ein Objekt mit einer bestimmten kanonischen Origo fest assoziiert ist, kann es auf der Basis der Stabilität der kanonischen Perzeptionsbedingungen seine Vorderseite sozusagen als objektzugehöriges Merkmal erwerben; das Grundprinzip ist aber auch hier die Perzeptionsperspektive der Origo."
>
> [The side of the object which points in the direction of the origin's line of sight (inside perspective) or the side of the object which is visible from the origin (outside perspective) receives the attribute 'front side'. The ascription of front is therefore a phenomenon that is *inferred* from the perceptual conditions of the origin. If an object is permanently associated with a certain canonical origin, it may, on the basis of the stability of the canonical perceptual conditions, acquire its front side, so to speak, as a feature belonging to the object itself; the underlying principle, however, is again the perceptual perspective of the origin.]

Thus, the front area of a car remains the same no matter whether the observer (origo) is actually sitting inside the car or not, since the position is canonical. It is viewed as the front side because the (canonical) person sitting inside the car can perceive objects in its front, which is not the case for the opposite direction. Notice that Grabowski's account is fully compatible with the previous observations in that perspective (outside vs. inside) plays a decisive role. The major difference is the influence of (canonical) perception and the view that it is not the lateral sides, but the frontal ones, which determine the two contrasting ways of axis ascription, quite contrary to the earlier term *handedness* used to describe this phenomenon.

Pederson (2003:291) notes that the interpretation of *clockwise* also depends on the perspective of an unspecified observer looking at a clock from the outside, while an inside perspective ("the perspective of someone lurking in the back of a grandfather clock") is easily conceivable, yielding a contradictory interpretation.

Determining an object's axes is thus often a matter of convention and conceptualisation with regard to the object's features and functionalities. But even in the case of abstract, non-functional objects speakers systematically view specific properties of objects as decisive for determining axes. For example, van der Zee & Eshuis (2003), who explore the influence of a relatum's three-dimensional properties on intrinsic usages of horizontal prepositions, find that long axes are preferredly taken as frontal axes, while the shortest axes were primarily used as lateral axes in (external) intrinsic reference relations. Relative curvature of the main plane of symmetry and contour expansion along a relatum's main axis also play a role in assigning reference axes.

2.2 Reference systems, topological distinctions, and perspective

As has often been noted, the same spatial situation can often be expressed in a variety of different ways. Perhaps the most impressive demonstration of this can be found in Talmy (2000:226), where the expression *in front of* can be interpreted in no less than four conflicting ways all of which are consistent with a certain spatial scene – and which, by far, do not exhaust the possible ways in which the spatial relationships present in that scene can be described. The reason for this is that these interpretations rely on the conceptualisation of different underlying reference systems. The aim of this section is to show in which ways reference systems can be employed, and point to the impact they have on the interpretation of dimensional terms.

Reference systems have been dealt with extensively in the literature, sometimes proposing that they belong to the semantic definition of dimensional expressions, sometimes dealing with them as a phenomenon entirely independent of language. Various different classifications and models have been proposed (e.g., Retz-Schmidt 1988, Herrmann 1990, Levelt 1996, Frank 1998, Eschenbach 1999), and the literature offers a variety of overviews and discussions of advantages and shortcomings of specific views (e.g., Levinson 1996, Pederson 2003). I will not continue this discussion but rather rely on one single approach that, to my mind, captures most of the distinctions highlighted in the literature so far, or at least is compatible with extensions capturing finer distinctions. Levinson (1996) elaborately describes the confusions arising with different and conflicting terminology, including various interpretations of terms like **deictic**, **extrinsic,** and **intrinsic**. Importantly, deixis is often confused with perspective, since both are based on the actual situation (cf. the discussion of accessibility presented in section 2.6 below). For this reason, in Levinson's alternative terminology (Levinson 1996, 2003) the term 'deictic' is avoided in the area of reference systems. Levinson's account, which is equivalent to Herrmann's (1990) "6-H-Modell", is the starting point I will adopt in the present work, and which I will extend with further insights stemming from other sources in order to propose a systematic comprehensive overview of possible reference systems in combination with the diverse perspectives available. To my knowledge, such an overview is not available in the literature so far, though many authors have pointed to many different aspects that need to be accounted for. All of the variants that I will present are relevant for the interpretation of dimensional terms; other spatial expressions need other kinds of frameworks (such as topological spaces) that will not be discussed here.

It should be noted that Levinson's main interests are much more far-reaching than simply providing a framework for reference systems. The work pursued in the "Space" research group which he is head of, part of the Language and Cognition department of the Max Planck Institute for Psycholinguistics (MPIP) in Nijmegen, aims at proving correlations between language and cognition based on detailed comparative research in many different cultures. With respect to reference systems, this may mean that (Levinson 2003:170):

> "[t]he frames of reference used in a language constrain or determine the frames of reference used by its speakers in the non-linguistic coding of spatial scenes."

This line of interest goes far beyond the scope of the present work where goals and evidence are overwhelmingly confined to the realm of linguistics, and moreover, focus on dimensional terms only. However, the fact that strong correlations with non-linguistic evidence exist proves that reference systems are not purely language-inherent phenomena.

2.2.1 External reference systems

Objects (or other entities) can be related either externally or internally; in the latter case one entity is located inside another entity. Levinson's account focuses on external relationships. It can be characterized schematically as recognizing three different reference systems with three variations each, dependent on whether the speaker, the hearer, or a third entity serves as the origin of the perspective employed. The three different options for reference systems are labeled by Levinson as **intrinsic**, **relative**, and **absolute**. They will be examined in the following subsections. The three options for origins reflect the distinction in grammar between first person (speaker), second person (addressee), or third person (cf. Herrmann 1990:131).

2.2.1.1 Intrinsic reference systems

In intrinsic reference systems, the relative position of one object (the **located object** or **referent**) to another (the **reference object** or **relatum**) is described by referring to the relatum's intrinsic properties such as *front* or *back*. Thus, in a scenario where a stone (the referent) is situated in front of a house (the relatum), the stone can be unambiguously identified by referring to the house's front as the origin of the reference system:

(13) The stone is in front of the house.

In such a situation, the speaker's and hearer's positions are irrelevant for the identification of the object. However, the speaker's or hearer's front or back, or, for that matter, left or right, may also serve as origins in intrinsic reference systems, as in:

(14) The stone is in front of me / you.

In this kind of intrinsic reference, the speaker's or hearer's actual position and orientation is decisive, which may be one reason why such sentences have sometimes been classified as deictic, causing the above-mentioned confusion in the intrinsic/deictic distinction.

In intrinsic reference systems, no further entity (such as, in our example, the house) is needed, which is why Herrmann (1990) refers to this option as **two-point localisation**. Put another way, the point of view or perspective, or in Levinson's terms: the **origin**, is conflated with the relatum, which is only possible if the relatum is at all capable of providing an intrinsic perspective, by virtue of its possessing internal parts. Eschenbach (1999:334) gives the following definition for intrinsic reference frames:

> "An *intrinsic relation R(F,G)* relates the positions of *F* and *G* on the basis of a frame of reference that is established by the functional-spatial structure of *G*."

Note that the "functional-spatial structure of *G*" must define the internal parts of *G*.

Intrinsic reference systems can also be induced by features of the figure rather than the ground, as already shown in example (12) above, repeated here as (15) for convenience:

(15) Sarah stood in front of the bush.

However, apparently only *in front of* is capable of expressing this kind of intrinsic relationship without implying an external observer, as the following example shows:

(16) Sarah stood to the right of the bush.

where an external perspective is needed for interpretation. Possibly, the relation expressed by (15) should be viewed as primarily functional rather than spatial, highlighting proximity and accessibility rather than a specific position on a spatial axis. But without an appropriate discourse context this remains speculative.

2.2.1.2 Relative reference systems

Humans employing relative reference systems, or, in Herrmann's terminology, **three-point localisation**, use the position of a third entity as origin instead of referring to inbuilt features of the relatum. Thus, the stone (the referent) may be situated to the left of the house (the relatum) from the speaker's, the hearer's, or a further entity's point of view (origin):

(17) From your / my point of view, the stone is to the left of the house.

(18) Viewed from the hut, the stone is to the left of the house.

Here, the house's front and back are irrelevant, which is why this reference system can be employed whenever the position of an object needs to be specified relative to an entity (a relatum) with no intrinsic directions, such as a box. Eschenbach's (1999) definition is as follows:

> "A *relative relation R(F,G,V)* relates the positions of *F* and *G* on the basis of a frame of reference that is established by the position and functional-spatial structure of an additional entity *V*. *V* may be implicitly given, it is different from both *F* and *G*, and it may also be a spatio-temporal process, such as a process of motion." [8]

In this definition, the fact that the view direction is derived from a certain position is expressed by assuming a further entity (such as speaker / hearer) with a "functional-spatial structure", such as a front, which permits assigning a viewpoint. This entity need not be expressed linguistically: thus, in (17) the phrase "from your / my point of view" could be omitted, which is often done in natural discourse (Herrmann & Grabowski 1994). Pederson (2003:290) proposes that, in English, it is possible to use a shortened version:

(19) The cat is to *your left* of the trash can.

which induces him to introduce a new term for this phenomenon, namely, **altercentric** or **hearer transposed relative** (as opposed to **altercentric intrinsic** as used, for instance, in example (14) above). Likewise, a **third party transposed relative** version is available, as exemplified by simply exchanging *his* for *your* in example (19). It could be added that it should be equally possible to say *to my left of the trash can*, using the speaker's perspective. According to the intuition of some native speakers of English, however, these options do not really seem natural and are usually dispreferred.

Notice that this abbreviated way of marking perspective is not available in German at all, which could be interpreted in the way that making the origin explicit in a relative reference system is not grammaticized in German. Furthermore, the options in English seem to be restricted (if they exist at all) to the lateral axis, as exemplified by the following:

(20) ? The cat is to *your front / back* of the trash can.

(21) The cat is (*to your) *behind / above / below* the trash can.

In example (18), contrary to Eschenbach's definition the point of view from a location is taken that itself does not provide a clear view direction, at least not if the hut's intrinsic properties are not known. Nevertheless, the spatial relationship can be interpreted uncontroversially (however vaguely), assuming that somebody (an entity possessing perceptual capabilities, which the hut itself does not possess) looks from somewhere close to (or inside of) the hut into the direction of the other two entities mentioned. Thus, although the description in (18) is underdetermined with respect to the exact position and view direction of the observer, and also with respect to the observer's identity and perceptual features, the spatial relationship may be sufficiently described for the purpose at hand – provided that it is given in a suitable situation. This is only the case if the hut (the origin) is located at some distance from the other objects, and not, for instance, between the two, yielding a situation in which there are two possible view directions toward the objects. Furthermore, the situation becomes more difficult if more objects are involved, so that it is not clear where the view direction should focus. In

[8] Eschenbach's observations concerning motion are taken up again in section 2.2.1.6 below.

specific situations this observation may turn out to be crucial, for example if the observer is not a person with easily identifiable perceptual abilities including view direction, but a robot with features that are hard to interpret or distinguish for a human participant.

If there are several objects of the same class present, it is furthermore possible to use the whole group of similar objects as a relatum. This option is called **group-based reference** by Moratz (e.g., Moratz & Fischer 2000). In contrast to situations where the relatum is an object of a different kind (here called **landmark-based relative reference system**), in a group-based relative reference system the relatum consists of one or several objects of the same kind as the target object. An example is the following:

(22) The ball is located to the left of the other balls (from my / your / a third point of view).

The term 'group-based' refers to the linguistic rather than a cognitive level. It does not aim at determining what kinds of factors (e.g., of distance or similarity) may lead people to conceive of an assembly of objects as a group, but rather, it simply points to the fact that people sometimes treat objects *linguistically* as a group, possibly based on perception, but maybe based on other factors as well. Such linguistic treatment is achieved by introducing all objects except for the target object as one (possibly plural) participant in a linguistic construction, as in *the other balls*. This participant can also be omitted on the linguistic surface while still being intended as relatum of the spatial description. In that case, the underlying reference system is in principle underdetermined, as the relatum could be the group of other objects as well as any other kind of relatum present in the scenario.

However, this is not necessarily true for dimensional adjectives. Eschenbach (2004) claims that German dimensional adjectives like *link-* (*left*) can only be employed when there is at least one other object of the same class present. She notes that they

> "cannot be used in predicative position (**Das Buch ist ober/vorder-* 'the book is upper/front'). They can be used only as restrictive modifiers to select an object (the figure) in contrast to another object (called here 'ground') of the same category. Thus, *das obere Buch* ('the upper book') refers to a book that is in a higher position than another book present. Nevertheless, in the context of a projective adjective the ground object cannot be specified by linguistic means but has to be provided by the situation of utterance."

Eschenbach thus formulates the intuition that projective (i.e., dimensional) adjectives require the presence of another object of the same class, as in (23) where the spatial relationship of one ball is determined (implicitly) by the position of at least one other ball present in the situation:

(23) Kannst du den linken Ball holen?
[lit.: Can you fetch the left ball?; "Can you fetch the ball on the left?"]

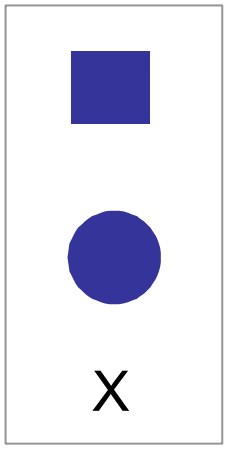
This utterance seems equivalent to (22) with respect to the localisation of the ball in relation to other balls present. Thus, it seems that in Eschenbach's definition the other object of the same class is used as a relatum, mirroring the present definition of a group-based reference system. Empirical analyses are needed to determine whether the employment of German (and maybe English, as well) dimensional adjectives is indeed restricted to cases of group-based reference, or whether other factors come into play as well. Note, however, that the literal English translation in (23) does not seem as appropriate as the German original. In English, intuitively a dimensional superlative (*leftmost*) would be preferred, or a completely different construction. Note that the English dimensional superlatives carry the restriction proposed by Eschenbach (the presence of other objects) in their semantics, since superlatives can only be used in comparison to other candidates.

Often, as seems to be true in this case, similar syntactic options are available in both languages but are not equally suitable in pragmatic terms. Therefore, although I try to keep translations close to the original syntactic format of the utterance, it should be kept in mind that this does not always yield an English utterance that could equally be employed in the

same situation – or vice versa – since English and German differ in their respective mappings of linguistic surface and spatial conceptualisations.

With regard to the frontal axes, Hill (1982) points to the fact that two conceptual alternatives are possible, one of which is used in languages such as English and German, and the other (at least) in Hausa. The two options are called **facing** versus **aligned**. In facing interpretations (as in English and German usage), an object that is located *in front of* another object is closer to the origin than the relatum is. In aligned interpretations (as in Hausa), the opposite is the case.

Figure 4. Two conceptual options for frontal axes in relative reference systems



In Figure 4, a facing interpretation would be that the circle is in front of the square from the point of view of X, while an aligned interpretation yields that the square is in front of the circle, as Hausa speakers would assume. Since the present work only deals with English and German usage, the aligned interpretation is simply viewed as an interesting conceptual alternative that is not put to use in these languages[9]. However, referring to the English and German usage as *facing* reflects the fact that many authors (e.g., Clark 1973) explain this kind of reference system as if the relatum was ascribed a "face" towards the viewer, regardless of whether it possesses intrinsic features or not. But the ascription of lateral and frontal axes with respect to the relatum is reversed: In surrounding space, axes are ascribed clockwise in the order *front-right-back-left.* In a facing interpretation with the relatum in the center, the order is *front-left-back-right.* For this (well-known) reason, the view adopted here is that the reference system originates in the viewer, not in the relatum; in this view the relatum is not ascribed a "face", but the viewer conceptualises regions from their own point of view that center around the relatum.

2.2.1.3 Absolute reference systems

In absolute reference systems, neither a third entity nor intrinsic features are used for reference. Instead, fixed-bearing systems such as *north* and *south* or, in some languages, *uphill* or *downhill* (Levinson 1996) serve as anchor directions. Thus, the stone may be located to the north of the speaker, the hearer, or the house. Eschenbach's definition for absolute reference frames is as follows (1999:334):

> "An *absolute relation R(F,G)* relates the positions of *F* and *G* on the basis of a frame of reference that is established by the functional-spatial structure of the common surrounding of *F* and *G*."

Pederson (2003:290) notes that there are at least two distinct subtypes:

[9] If the entities involved are part of a queue, then interpretation would be based on the functional front end of the queue rather than X's point of view, which would then lead to similar results as in the Hausa example (i.e., the square is in front of the circle with respect to the queue). Similar effects arise if the entities are in motion (see section 2.2.1.6), where the direction of motion induces a perspective on the situation.

> "1) an absolute RF which relies on co-ordinates which do not derive from the immediate environment and 2) an absolute RF which creates ad hoc co-ordinates by appeal to perceptual features."

An example for the first option is the usage of the earth's cardinal directions (north, south, east, west), and for the second, the usage of local landmarks for projective descriptions (Pederson 2003:290):

(24) The cat is towards the wall from the trashcan.

Absolute reference systems are a special case in that there is no way of labelling origins or relata in a way consistent with the other kinds of reference systems, as directions behave differently than entities. Pederson attempts to capture the relationship between his second kind of absolute reference systems and intrinsic reference systems as follows (2003:290):

> "In both cases there is a direction determined from the ground on the basis of perceptually available features. The difference being that in one case the features belong to the ground and in the other case they do not. It is unclear whether to treat *towards the wall* as a case of transposing co-ordinates onto the trashcan. Subjectively, the example seems to represent a route from trashcan to wall, along which a cat will be found. As such, *the wall* is essentially the same indicator of direction as a more abstract term like *North*."

It seems that spontaneously created absolute reference systems (Pederson's second kind) are not very well researched so far. Absolute reference systems of the first kind are, in Western societies, only used in specific cases, preferably in large-scale space, as opposed to indoor scenarios (which are mostly in focus in the present work); and preferably in relation to maps (Levinson 2003). Crucially, intrinsic and relative reference systems are used with dimensional expressions, while absolute reference systems (of both kinds) employ a completely different linguistic repertory. For these reasons, they will not be analysed specifically in the present work.

It should be noted, however, that the terms used for absolute reference systems may also be used for other kinds of spatial reference in certain contexts. For instance, resulting from the fact that maps are oriented with respect to the absolute directions, placing North at the top, it might happen that the top side of a depiction is referred to as North even if this is not actually the case. Furthermore, Carroll (1997) reports that native speakers of English, in an experimental study involving models of villages and rooms, used *north, south, east, west* in a non-absolute way (1997:140):

> "The use of the 'earth' axes in English is apparently warranted by the fact that rooms and villages are places we can walk around in. The axes are not absolute but are established relative to the position of the observer, so that the north side is typically the side furthest away from the viewing point."

This usage corresponds to the usage of dimensional terms in internal regions in the visual field of an observer (see section 2.2.2.2 below).

2.2.1.4 Choice and coincidence of intrinsic, relative, and absolute reference systems

Table 3 exemplifies usages on the horizontal plane in the two reference systems in focus, presenting the impact of variability of the origin on the language used.

Carlson (1999:367) points out that, on the vertical axis, gravity defines an absolute reference frame that is often, but need not be, coincident with intrinsic and relative interpretations:

> "Sometimes the axes of the three reference frames are aligned, such that the orientation and direction parameters are set similarly. For example, gravity, an upright viewer, and an upright reference object all indicate the same direction for the spatial term 'above' (…). However, because people and objects in the world can move and rotate in space, sometimes the axes of the three reference frames are misaligned and assign competing directions to the same spatial term."

This observation holds also for the horizontal plane, except that here, it is impossible to align relative, intrinsic, and absolute reference frames in any consistent or generalisable way. This is due to the fact that gravity as source for an absolute reference frame coincides with the prototypical vertical direction (cf. section 2.1), while horizontal directions are usually derived from view directions and internal object parts, both of which are completely independent of absolute directions. As Levinson (2003:35) puts it:

> "[P]erceptual cues for the vertical may not always coincide, but they overwhelmingly converge, giving us a good universal solution to one axis. But the two horizontal coordinates are up for grabs: there simply is no corresponding force like gravity on the horizontal. Consequently there is no simple solution to the description of horizontal spatial patterns".

Table 3. Linguistic examples for combinations of reference systems and perspectives

origin	**relative**	**intrinsic**
speaker (1st person)	The ball is in front of the tree from my point of view.	The ball is in front of me.
listener (2nd person)	The ball is in front of the tree from your point of view.	The ball is in front of you.
third entity (3rd person)	Viewed from the stone, the ball is in front of the tree.	The ball is in front of the arrow.

Put in more concrete terms, objects can be conceptualised as having intrinsic properties that define all three axes, such as a car whose bottom can be defined by its wheels and the top by its roof, the front by its prototypical motion direction, and the lateral sides by the positions of the seats for people to sit in. These directions remain true even if the car is turned upside down, in which case two possible ways of talking about the car's top can be derived – the intrinsic top and the current upper side. However, this applies mainly for internal relations (see section 2.2.2.1 below) but less straightforwardly for external ones, as research by Carlson (1999) has shown. Her study concerns the activation of different reference frames with respect to the vertical axis, i.e., the assignment of an *above* relation on the basis of either the object's features or gravity (viewed as an absolute reference frame), which was clearly preferred. She concludes (1999:376):

> "The research presented in this paper focused on the process by which a reference frame is selected. Together, the results indicate that reference frame selection involves the simultaneous and automatic activation of multiple reference frames, followed by a selection process that is assisted by inhibition of the non-selected frames. (…) [D]ifferent reference frames may be used online during apprehension because people have different preferences for using these reference frames to define various spatial terms."

She also notes that, on the horizontal plane, preferences may well be different, possibly favoring relative reference frames for the front region, and reflects on the relation to non-linguistic reference frame selection as addressed in several studies.

In describing non-intrinsic *above* relations, Carlson (1999) conflates absolute and relative reference systems, which, in her work, makes sense because viewers were always in canonical position. Somewhat similarly, Levinson (2003) (who generally focuses on the horizontal axes) claims that the vertical dimension "shares the same major divisions into intrinsic, absolute, and relative" (p66). But it is questionable whether, in cases where the observer is not in upright position, relative reference systems can be derived at all. In that case, the position of a viewer would need to externally determine a vertical direction that is independent of both gravity and the object's internal properties. However, parallel to the examples listed in Table 3 above, it does not seem to make sense to say:

(25) * The ball is above the tree from my point of view.

in a situation where the speaker is not in upright position, and the ball not in a higher position than the tree (in relation to the earth). Or to use a different example: Imagine a person in the air above a swimming pool who has just jumped head-first from a height of five meters. From the perspective of this person, the pool's ground floor is *above* the water – but there is no straightforward way the person could actually put this conception into words and be understood. He could, of course, verbalise the perception as such, as in:

(26) From this position, it really looks like the pool's bottom was above the water.

which is somewhat different from stating that the pool's bottom *is* above the water from someone's point of view.

One study that directly addressed this issue is presented by Carlson-Radvansky & Irwin (1993). They attempted to trigger a relative reference frame for *above* by placing the viewer in a lying position and showing them pictures. The referent (a point-like fly) appeared in a roughly aligned position with the observer's head next to the relatum (a chair). Results showed no independent contribution of the relative reference frame: In the described situation, *above* was only used if the chair was also in a lying position, in which case an intrinsic reference frame was activated. However, *above* was clearly more favoured in situations fitting the absolute (gravitational) reference frame.

According to research by Friederici & Levelt (1990), people use their own head position (i.e., their perspective) for reference in the absence of gravitational cues. The subjects were asked to describe the position of a white ball with respect to a black ball. Since balls do not possess intrinsic axes, intrinsic reference systems could not be employed. As gravity was not available for reference, the subjects could only define the vertical direction by using either their head position or the visual background (pictures of oriented trees) as a cue. The nature of this task suggests to me that the reference system employed was not genuinely relative in the sense that the subjects could have said "from my point of view". Instead, they employed an absolute reference system based on an assessment of the main axis by cues other than the normally available gravity (see section 2.2.4 below). Thus, either their own head position or the visual background served to define where the vertical axis could be, in order to define the absolute directions. Other kinds of tasks would be necessary to determine whether relative reference systems are available with the vertical axis in the same way as with the horizontal axes: for example, providing different possible points of view that conflict with respect to their "vertical" direction. The main aim and outcome of Friederici & Levelt's study was not the proof of an existence of relative reference systems with the vertical axis, but rather, to find out how people define the vertical in the absence of gravity. Thus, it is a remarkable result that (1990:265):

> "although on earth gravity plays a dominant role for the choice of spatial reference in adults, mental representations or concepts of spatial orientation can be used quite consistently in its absence. (…) When gravitational cues are absent or task-irrelevant, subjects tend to use the head-retinal reference system."

Pointing to the controversy with regard to whether terms denoting the vertical axis should be included in the group of projective terms (i.e., those terms that need to be interpreted on the basis of a reference system), Retz-Schmidt (1988:97) summarises the suggestions of several authors, showing that some major approaches only deal with horizontal terms (*in front of, behind, to the left of, to the right of*). She then points out that "deictic" (in our terms, relative) uses are not possible with the vertical axis, though "extrinsic" (absolute, gravity-based) and intrinsic interpretations are.

Thus, with the vertical axis, the preferred interpretation is absolute rather than intrinsic, and a relative interpretation is (at least nearly) impossible. What about the horizontal axes? Clearly, in this case absolute reference systems play a minor role in Western societies, as opposed to some other cultures where they are used throughout, and where relative reference systems

sometimes are not even available at all (Levinson 2003). Miller and Johnson-Laird (1976:399) claim that people initially assume *intrinsic* interpretations:

> "In interpreting spatial indications, people first determine whether the landmark has intrinsic parts. If it does, they try to interpret the spatial relation intrinsically unless they are explicitly informed to the contrary. If the landmark does not have intrinsic parts relevant to the spatial indication, they must rely on context to provide a deictic interpretation. If both strategies fail, they may ask for more explicit information."

Similarly, Ehrich (1985) assumes that intrinsic usages are preferred in contrast to "deictic" (relative) ones. Levinson (2003:108) observes that the expressions used for relative reference systems are usually based on those for intrinsic ones:

> "Normally, then, the relative frame of reference is expressed through the same classes of expressions as the intrinsic ones, from which they are normally derived by diachronic extension."

But with regard to usage, the MPIP research yields a rather different picture (Levinson 2003:179). In contrast to languages that only offer one kind of reference system for a situation,

> "for example in English or Dutch, both relative and intrinsic frames of reference are available and colloquially used, but the relative frame is clearly predominant for most kinds of spatial description. In the case where more than one frame of reference is available, one may find one frame of reference preferred for one situation, and another for another situation."

Since the main aim of research at the MPIP is headed another way, this situation-dependency is acknowledged but not investigated systematically across languages. However, the valuable insight is recognised that only similar kinds of tasks can be compared effectively between cultures.

Note, furthermore, that the question of variation between reference frames entails several different aspects. Some research (e.g. the work by Carlson 1999 described above, and other contributions by the Carlson-Radvansky research group) deals with the question of how reference frames are processed, whether one "default" interpretation is activated first, if all possibilities are active at all, and if yes, if that is a serial or a simultaneous process. Taylor et al. 1999 give an overview of research in this direction and apply a technique of cognitive neuroscience (event-related potentials, ERPs) to this question. Comparing processing of intrinsic versus relative reference frames on the horizontal plane, they found a priority selection of the intrinsic reference frame. The relative frame was only activated later, i.e., retrieved from memory if needed for interpretation.

The etymological and structural features of the terms applicable for the diverse reference frames as indicated above represent a different direction of research. But the question that is most relevant to present purposes is the preference of reference frames in application rather than cognitive processing, ideally dependent on cultural or language-specific as well as situational factors. In this direction, very few results are available so far. Carroll (1997) describes differences between English and German usage of reference frames. Unfortunately, she uses a classification (a binary distinction between intrinsic vs. deictic involving partly perspective, partly reference systems) that cannot be directly mapped to the present approach; furthermore, she seems to presuppose a direct correspondence between specific syntactic forms of dimensional terms and reference systems that is not taken for granted in the present approach. Accordingly, it is in many cases not transparent how her classification of usages should be interpreted and related to the present classification. The main result seems to be that German speakers tend to rely on their own perspective, while English speakers preferably use the intrinsic properties of the objects involved (in which case speaker perspective is irrelevant).

Another study pointing to differences between German and English usage is presented by Weiß et al. (1996), who investigated speaker's choices between intrinsic and relative reference

systems in a car driving situation in which a parking lot should be identified. The two possible parking lots were situated at the two frontal sides of a parked car or, alternatively, a tree. Both parking lots could be referred to using either the front or the back half-axis, depending on whether the speaker used an intrinsic or a relative reference system. Thus, an utterance like "in front of the car" was ambiguous because it could either refer to the parking lot in front of the car's front side, or alternatively to the parking lot between the observer (approaching in another car) and the parked car. By exchanging the parked car (which offers an intrinsic reference system by its internal sides) for a tree (which does not possess internal sides), the impact of the presence of an oriented entity as relatum could be investigated. It turned out that the frontal dimensional terms are highly ambiguous in both languages, both in production and in comprehension. A probability computation of the results yielded a speaker/hearer correspondence in only slightly more than 60% of cases in German and 70% in English. The spatial dimensional terms were nevertheless used spontaneously and without further explication by more than half of the participants in the production task. The experimental participants generally did not seem to be aware of the underspecification of the description. The most interesting differences between German and English concern the usage of *vor* in German, which can also be interpreted temporally (in the sense of *before*), and the choice of intrinsic versus relative reference systems. It is possible that the seemingly relative usages of *vor* were intended to denote temporal priority rather than a spatial position based on the speaker's viewpoint relative to a relatum. However, *hinter* (which is only spatial) was also used in ambiguous ways, so that this can only explain part of the findings. In English, whenever there was a relatum with intrinsic features available, it was used consistently for reference. This finding is in line with earlier assumptions in the literature (e.g., Miller & Johnson-Laird 1976) that speakers prefer intrinsic interpretations. Ambiguity only arose when the relatum was not oriented (i.e., in the case of a tree serving as relatum). In such cases, speakers preferred other kinds of expressions in the production task. In German, in contrast, orientedness did not seem to play a major role. Instead, there was a preference to refer to the parking lot that was closer to the driving car, regardless of the preposition used. Furthermore, an influence of the discourse situation could be noted: the German speakers used intrinsic reference more often in a driving-test situation than in a more informal (drive-home) situation. Recent research by Vorwerg & Weiß (2004) shows that this difference can be traced back to the exact wording, specifically, the verb used in the instruction to the participant. In German, the driving-test situation involved the expression *einparken* ("Parken Sie bitte vor/hinter dem gelben Käfer ein", i.e., "Please park in front of/behind the yellow beetle"), while the drive-home situation involved *rauslassen* ("Laß mich doch bitte vor/hinter dem gelben Käfer raus"; i.e., "Please let me out in front of/behind the yellow beetle"). This new finding highlights the influence of other lexical items in the discourse context on the interpretation of locative dimensional terms.

It should be noted that the situation investigated by Weiß et al. is highly specific in several respects. First, it does not take place in the real world, but rather, the participants are presented with a drawing on a cardboard, and a toy car standing on the drawing. They are then asked to imagine driving the car. Thus, they are presented with a static situation but asked to imagine a dynamic one. The impact of static versus dynamic starting positions, however, is not addressed, in spite of the fact that motion induces a perspective of its own (cf. section 2.2.1.6 below). Note, furthermore, that the temporal interpretation of German *vor* is specific to the dynamic situation. Second, the imagined traffic situation suggests a necessity for a quick decision, one which, furthermore, is unlikely to have consequences of any kind even in the case of misinterpretation. It would have been interesting to see what happens if the parking lot decision had been ascribed specific importance. The formal driving-test situation (which is indeed shown to have consequences at least for the German speakers' choices of reference frames) may suggest some such importance, but only in an implicit fashion.

Because of these special circumstances, the generalisability of the results is not clear. It stands to reason that other kinds of situations offer more (or other kinds of) clues for the decision between reference frames. Thus, the influence of contexts and discourse tasks (section 5.1) should not be underestimated.

One study by Carlson-Radvansky & Radvansky (1996) also points in this direction. Unlike most studies in Carlson's research group, this study deals not only with the vertical axis, but investigates the impact of a functional relationship (cf. section 2.8) between the referent and relatum on the choice of intrinsic versus relative reference systems with all three dimensions. It turns out that a functional relationship, such as a mail carrier who is oriented towards a mailbox, enhances acceptability and production of dimensional terms in an intrinsic reference frame. Where such a functional relationship is ruled out, e.g., if the mail carrier is oriented with his back to the mailbox, relative usages become more likely. Unfortunately, this study does not inform about any kinds of differences between the axes, which are likely to occur considering the findings presented in sections 2.1 and 2.8. Moreover, a production task shows that subjects to a very high degree choose other kinds of reference frames which are not explained in the paper, but merely subsumed in a large category "Other", which in the non-functional case covers as much as 50% of the results. Nevertheless, the finding that the choice of an intrinsic reference system rather than a relative one is influenced by the presence of a functional relationship is plausible because "intrinsic terms focus attention on the object" (Carlson-Radvansky & Radvansky 1996:56) in the sense that intrinsic sides can only be attributed when considering the object's properties (cf. section 2.2.1.1). These are often determined by the object's functional features (if not by perceptual ones, as in the case of animated beings).

Two studies following up this research direction were presented by Taylor et al. (2000, 2001). They investigate, on the one hand, speakers' choices with regard to either intrinsic or relative ("deictic" in their terms) reference frames, and on the other hand, the choices of the relatum versus referent (cf. section 5.2.2). Participants were presented with drawings of two objects and were asked to describe their spatial relationship using a sentence frame such as "The _____ is _____ the _____". Results show that there was no general preference for either kind of reference frame. With familiar objects, an intrinisic reference frame was chosen more often for functionally oriented objects and a relative frame for non-functionally oriented objects (Taylor et al. 2000). However, with novel objects this result was not replicated, although the participants were made aware of the functional relationship in advance (Taylor et al. 2001). In this task, there was a tendency for speakers to stick to one strategy across trials.

Ehrich (1985) notes some interesting empirical results about descriptions of the interior of a room: Speakers seem to take into account that some reference options involve possible conflict with other options, while others do not. Where conflicting interpretations are possible – mainly, with regard to the lateral axis – speakers prefer other kinds of descriptions such as temporal order (*dann, danach*) or use topological or direction-neutral expressions such as *an, bei* or *neben.* Furthermore, she observes that sometimes a "deictic perspective" (which, presumably, is to be interpreted as equivalent to a relative reference system in this case) is preferred if this can guarantee a constant point of view from which a scenario is described. With intrinsic reference systems, the perspective would have to be changed for each relatum to be described, which would make the description much more complex. An intrinsic reference system can obviously only be used if the relatum possesses internal axes, which makes a constant reliance on intrinsic systems unlikely. In the experiment described by Ehrich, perspective is held constant in 95% of cases (Ehrich 1985:159). Note, however, that speakers in many other cases do switch perspectives, as described in section 5.2.1 below.

Apart from these few results, unfortunately, there has not been much research to determine what kinds of reference system are preferred in linguistic tasks dependent on what kinds of

factors. This is not surprising considering the fact that comprehensive systematic accounts of reference systems were first presented during the 1990's. Furthermore, decisions between landmark-based and group-based relative reference systems cannot be predicted at all since the latter kind of reference has not been recognized widely at all so far.

On a wider scale (encompassing several of the topics addressed in later sections of this report), research results provide some insights into speakers' choices of strategy in spatial descriptions. These will be addressed specifically in section 5.2.1 below. It turns out in this line of research that ease of reference as well as discriminability play a role in such strategy choices, a finding that certainly applies also to the narrower area of choice of (external) reference systems. Since spatial axes are cognitively not equal, this may mean that a reference system is chosen that avoids the lateral axis in favour of the frontal one (Tversky et al. 1999). Furthermore, processes of interactive alignment (Pickering and Garrod 2004) lead to a preference for a kind of reference system that has already been employed by the interlocutor (Watson et al. 2004), even if this entails choosing a different spatial preposition (contrary to the effects of lexical priming) (see also section 5.3.1 below).

2.2.1.5 Choice and allocation of perspective

With respect to the perspective employed, more results are available in the literature – which is also not surprising since the impact of the underlying perspective on the interpretation of the spatial description has been acknowledged in the literature for decades. According to Herrmann & Grabowski (1994), speakers employ the listener's point of view specifically if there are reasons for this. Obviously, processing is easier for the listener if their own perspective has been used; thus, if the speaker wishes to facilitate interpretation for the interlocutor, they use the interlocutor's viewpoint. This is especially likely if the partner is expected to have less cognitive or linguistic capabilities, e.g., because of younger age or because someone is not a native speaker of the language employed. But it can also be the case in situations where the speaker wants to be polite, e.g., if the interlocutor has a higher status. Furthermore, if actions are involved on the part of the listener, usually the listener's perspective will be used.

Tversky (1999) reports similar results concerning the social implications of adopting another person's spatial perspective, concluding that there is no default perspective; rather, adopting different perspectives on one occasion is not uncommon. In contrast, Herrmann & Grabowski (1994:123) propose that, in the absence of specific reasons for taking the interlocutor's perspective, the listener should assume that the speaker uses their own perspective as a standard strategy. This is because specific mental costs are required for the adaptation of another person's perspective. As a long tradition in psychology (starting with Piaget's extensive research, e.g., Piaget & Inhelder 1956) has shown, it takes children several years to be able to mentally adopt a different perspective than their actual one. These costs may differ in situations involving different degrees of divergence from the interlocutor's view direction, and they are not primarily a problem of language (Herrmann & Grabowski 1994:142):

> "Nach den gegenwärtig verfügbaren Befunden ist der Sprecher deshalb in der Vis-à-vis-Situation nur mit Anstrengung in der Lage, partnerbezogen zu lokalisieren, weil mentale Selbstrotationen in diesem Bereich großer Winkel *generell* erschwert sind. Die relative Leichtigkeit der partnerbezogenen Lokalisation in mittleren Winkelbereichen kann wahrscheinlich darauf zurückgeführt werden, dass die mentale Selbstrotation im Bereich des Greifens und Manipulierens *generell* leicht vonstatten geht. Insofern ist das Problem der mentalen Kosten für partnerbezogenes Lokalisieren in erster Linie kein sprachliches Problem."
>
> [According to the currently available results, therefore, in face-to-face situations speakers require much effort for localising in a partner-oriented way, because mental self rotations are *generally* difficult with large angles. The relative ease of partner-oriented localisation with medium angular ranges may possibly be related to the fact that mental self rotation is *generally*

easy within grasping and manipulation areas. Therefore, the problem of mental costs for partner-oriented localisation is not a specifically linguistic problem.]

The impact of different angles in ego-reorientation on the mental costs involved has been studied extensively by v. Wolff (2001), who shows that egocentric axes can be decisive as well as the symmetry axes of a room, depending on the circumstances. Angles diverging from these basic axes are harder to adopt.

These results show that the cognitive load of adopting a different person's perspective is indeed higher than using one's own point of view. Accordingly, it can be considered more polite to take over the cognitive load, and it is plausible that this is done for specific kinds of addressees and dependent on the scenario. Further evidence for this hypothesis stems from research by Schober (1993,1998) and Mainwaring et al. (2003). Schober's findings prove that, with real interaction partners, the proportion of speakers using their addressee's perspective is much higher than with imagined interaction partners. Mainwaring et al. show that the addressee's perspective is preferred across a range of different configurations and scenarios, except for two cases: The speakers' perspective is preferred in tasks where, although there is a further (imagined) participant, the addressee is the speaker's own future self rather than the other person. Furthermore, in situations where the target object was situated closer to the speaker than to the addressee, speakers tended to refer to their own rather their addressee's position, using distance terms rather than dimensional terms. In the present framework, this observation does not really (contrary to Mainwaring et al.'s interpretation) relate to perspective choice but rather to choice of linguistic strategy, since dimensional terms differ in certain fundamental respects from distance expressions (see section 5.2.1 below). Crucially, they do not depend on a view direction at all, because distance terms only require a relatum, not an origin (of a reference system).

Another finding worth mentioning from Mainwaring et al.'s study is a comparison between a task in which the addressee was to be *informed* about the target object's location, versus a task in which the addressee was to be *asked* about the target object's location via a yes/no question (see also section 5.1 below). While the addressee's perspective is preferred in both tasks, the proportion of utterances using the speaker's perspective is considerably higher in the question task. The authors explain this finding by the assumption that the cognitive load is more on the speaker in the question task, because the addressee possesses the relevant information.

With regard to the third entity perspective, Herrmann (1990) notes that it is unlikely to be used in indoor scenarios. His example (1990:129) indeed sounds strange:

(27) ? Vom Stuhl aus gesehen steht die Vase vor dir.
[Viewed from the chair the vase is in front of you.]

This is particularly deviant because, in this case, the addressee is used as relatum (rather than origin), ignoring the addressee's orientation, and an inanimate object lacking perceptual abilities is used as origin. Nevertheless, Herrmann's suggestion that third entity perspectives are especially likely to occur in non-face-to-face situations is intuitively appealing (1990:128):

"Ihre Hauptdomäne ist wohl die vorstellungsmäßige Hineinversetzung des Hörers in eine aktuell *abwesende* Raumkonstellation (…). Äußert S zu H: 'Vom Bahnhof aus gesehen liegt das Hotel direkt vor dem Wasserturm.', so mag der Bahnhof das 'vorzustellende' (de facto abwesende) Objekt O sein, das die Origo besetzt."

[Their main domain is probably the imagined positioning of the listener into a currently *non-present* spatial configuration (…). If S says to H: 'Viewed from the railway station, the hotel is located directly in front of the water tower.', then the railway station may be the 'to be imagined' (in fact non-present) object O, which the origin occupies.]

Furthermore, there may be situations in which it is useful to ignore the listener's actual view direction since it may be presently unknown (Herrmann 1990:128f.):

"Falls zum Beispiel der Sprecher unterstellt, er kenne zwar (…) die *Position* eines Menschen im Raum, aber nicht dessen aktuelle *Raumrichtung*, so kann er dem Hörer doch eindeutige Informationen der folgenden Art geben: 'Von der Scheune aus liegt der Schlüssel genau hinter Otto'."

[If, for example, the speaker implies that he knows (…) the spatial *position* of a person but not the person's current *spatial orientation*, he can still give the listener unambiguous information of the following kind: 'Viewed from the barn, the key is located directly behind Otto.']

Other results in the literature concerning choice of perspective in linguistic interaction mostly pertain to discourse-related factors, such as consistency within a discourse and co-ordination with the interaction partner (see section 5 below).

Specifically in route descriptions, the speaker can **induce** the interaction partner to orient themselves in a specific direction, as in (Herrmann & Grabowski 1994:133):

(28) Drehen Sie sich doch einmal um und stellen Sie sich so hin, dass Sie da hinten das chinesische Restaurant sehen.

[Please turn around and position yourself in a way that you can see the Chinese restaurant at some distance.]

Kray (2003) has explored this option for application in a mobile tourist guide, a handheld device to be used for navigation through unfamiliar surroundings. He defines as follows:

"An *induced frame of reference* is a frame of reference that requires the listener to first perform one or more mental or physical actions before the frame of reference is established. These actions include rotation and relocation, which may be applied to the origin and/or the orientation of an original frame of reference, as well as defining the handedness."

For instance, in contrast to Herrmann & Grabowski's example the device may not know the actual orientation of the listener and can therefore not simply suggest rotation as in example (28). Instead, it may suggest a view direction on the basis of nearby buildings:

(29) If you stood in front of the church, the fountain would be to your right.

2.2.1.6 Perspective induced by motion

Objects in motion (or potentially in motion) induce a further kind of perspective that is independent of intrinsic fronts or perceptual organs. The perspective adopted can be described as though the moving object was viewed from the inside, so to speak, looking in the direction of motion. Thus, even completely symmetric objects, such as a ball, can be ascribed front, back, right and left sides when in motion:

(30) John is running to the right of the ball that is rolling down the hill.

(30) can also be interpreted with a relative reference system using an external viewpoint, so that the utterance is potentially ambiguous. It is not trivial to predict which of these options would be preferred in natural discourse.

Furthermore, the assignment of sides in the case of motion holds for scenarios in which the moving entity itself serves as the origin of a relative reference system. In this case the perceptual apparatus does not play a role, but rather the direction of motion, as in:

(31) If the ball continues rolling down the hill, it will fall into the hole in front of the wall.

This observation leads Eschenbach (1999:334) to include motion as a possible origin of reference systems in her definition of relative reference frames (see p 22 above). She notes (Eschenbach 1999:341) that, in this case, the relation

"is judged with respect to a third moving object *X* (*die Steigung ist vor dem Ziel* / the slope is in front of the finish). This shows that motion based uses of *vor* and *hinter* consider motion of *X* relative to *G* (…) rather than motion of *G* or *F*."

In the present approach, the effect induced by motion is not directly related to reference systems but rather to perspective. Thus, any oriented object can serve as the origin of both intrinsic and relative reference systems; the orientation can come about by (potential) motion

or by intrinsic features such as perception. Furthermore, similar effects arise by functional ordering relations such as those induced by a queue (see section 2.1 above).

2.2.1.7 Concepts of perspective

It is often pointed out that perspective involves more than just the point of view of an observer. Langacker (1986:12), for example, states that this notion

> "subsumes a number of more specific factors: orientation, assumed vantage point, directionality, and how objectively an entity is construed."

In this terminology, "assumed vantage point" corresponds to the view direction used (in the present classification) to specify reference systems. Additionally, the orientation of the observer or an object whose internal parts are relevant for a spatial description can be important. If the orientation of the observer is not clear they can still be motivated to adopt a specific view direction, e.g., by referring to a certain assembly of objects that are located in the vicinity of the observer. The following utterance combines a reference to the observer's actual orientation and then to the intended vantage point:

(32) There are three objects behind you. The leftmost of those is the one you need.

The reference can be resolved without perception, i.e., without turning round. Thus, referring to objects can induce discourse participants to adopt a certain vantage point – actually or mentally – independent of their actual orientation even in a face-to-face situation.

Directionality, in Langacker's terminology, concerns the conceptualisation of a situation as directed from a certain source. This may be the case even if the situation does not involve motion, as in his examples (1986:13):

(33) The hill falls gently to the bank of the river.

(34) The hill rises gently from the bank of the river.

in which some kind of abstract motion is implied, which may involve an underlying implicit conceptualisation of a potential traveller. In contrast, if moving objects are described, then situations are construed in a more "objective" way, i.e., differentiated from the conceptualiser, as in

(35) The balloon rose swiftly[10].

Such gradations of physical versus abstract motion, and objective versus subjective construals, though spatial in nature, will usually not play a role in interpreting reference systems so that they are simply mentioned here for the sake of clarity and differentiation. In more general terms, these distinctions show that different kinds of perspective always come into play when using language. All utterances in some way reflect the speaker's position with regard to many different dimensions, few of which have direct consequences on the interpretation of dimensional terms. Most important for present purposes is the distinction between actual orientation and vantage point, which are often conflated because they are taken to coincide, which in fact they often do.

Langacker's distinctions, however, do not cover all (spatial) senses in which the term 'perspective' is used in the literature. Another sense emerges in work on different kinds of knowledge about spatial relations, where the well-known difference between **route** and **survey** representations is often expressed in terms of perspective (e.g., Taylor & Naylor 2002). In the area of language, these distinctions are especially relevant with respect to speakers' strategies in describing a spatial situation (see section 5.2.1 below). See also Tversky (1999) for further reflections on the notion of (spatial) perspective.

In the present work, the term perspective is used solely for the origin (vantage point) of a reference system as used in spatial descriptions.

---

[10] Note that the interpretation of *swiftly* naturally depends on the conceptualiser's subjective assessment.

2.2.2 Internal relations

The discussion so far has concerned solely **external** relationships, i.e., relations between objects that are spatially separate. These need to be contrasted with **internal** relationships, which depend either on the conceptualisation of regions inside objects, i.e., object parts (section 2.2.2.1), or on the partitioning of a region, such as the observer's visual field (section 2.2.2.2).

2.2.2.1 Object parts

We have already seen that, for the application of intrinsic (external) reference systems, the origin / relatum needs to have internal properties, such as a front. For the description of these object parts, the same dimensional expressions (*right, left, front, back side*) are employed as for relationships between objects. Furthermore, other objects can be located inside of other objects and then be described on the basis of the relatum's internal parts:

(36) The table is standing in the back of the room.

(37) Der Tisch ist hinten im Zimmer.

(38) The woman is sitting in the front of the car.

(39) Die Frau sitzt vorne im Auto.

Note that a typical English surface form for such a relationship is a noun in a prepositional phrase using *in*, while in German the adverb is used (see also section 2.2.3 below).

How objects' internal parts are derived was dealt with in section 2.1. Apart from the processes described there which pertain to permanent object properties, internal object parts can be ascribed spontaneously and temporarily, as noted by Herrmann (1990:121):

> "Beim neben dem Waldweg stehenden Baumstumpf ist diejenige Seite 'vorn', die der Stumpf dem auf dem Wege wandernden Beobachter zuwendet."
>
> [The front side of the tree stump that is situated next to the forest track is the side that points toward the observer walking on the track.]

Thus, it is possible to interpret an utterance like (40) even though trees are usually not ascribed internal parts, lacking typical perceptual organs, functionally salient sides, or direction of typical motion.

(40) The front side of this tree really looks scrubby.

Here, the front side of the tree is not ascribed on the grounds of typical features, but *ad hoc*, viewed from the outside (cf. section 2.1). There may be various reasons for ascribing sides in an *ad hoc* way. In this case, if (40) is uttered in a similar context as that described by Herrmann (with respect to a stump), it is probably the observer's visual access to the tree that is responsible for the spontaneous ascription.

Conceivably, it may also be possible to ascribe an *ad hoc* inside view interpretation (Herrmann 1990:139), for instance, if someone is presently occupying the object:

(41) The left side of the tree in which you are sitting really looks scrubby.

Here, the left side of the tree might be determined by the addressee's position, for example, if the speaker's own view direction is less clear. However, such a situation seems far-fetched; it is probably usually the onlooker who determines the *ad hoc* ascription of sides.

Object parts can also be ascribed via the perspective induced by motion (cf. section 2.2.1.6), as in the following example in which the entity (an army) is partitioned into sections:

(42) The front section of the army moves much too fast.

2.2.2.2 Partitioning a region into internal parts

Regions can be partitioned into internal (relative) sections by adopting a "global perspective" (Carroll 1993, 1997) from an observer's viewpoint (see also section 5.2.1 below). In that case, the observed region is divided into spatial sections and ascribed part regions that are described

by the dimensional terms (*front, back, left, right)*, sometimes explicitly so by referring to *sides* (such as "on the left/right side", Carroll 1993:30). The observed region can be a specific assembly of objects that are perceived as belonging together or being relevant for the discourse situation, or any other kind of region that is within the limits of perception. For example, in German it is possible to say:

(43) Siehst du den Apfelbaum dort hinten?
[lit., "Do you see the apple tree there in the back?"]

where the visual field is partitioned into regions in relation to the position of the speaker (and, in a felicitous context of the utterance, the hearer as well). Then, the area close to the observer is referred to as *vorne* (front), and the area more distant to the speaker within the visual field is referred to as *hinten* (back). Note that, in the case of *front, left, right* the regions coincide with the intrinsic perspective of the speaker, but not in the case of *back*, because the intrinsic back region is located behind the speaker, while in the case of partitioning the visual field the back region is a region in front of, but distant from the speaker. Furthermore, no further objects need to be involved that could provide the origin or relatum of relative or intrinsic reference systems. Thus, utterance (43) is based on a different kind of reference system, namely, the partitioning of a certain region (e.g., the visual field) into (internal) subregions of that region.

Carroll (1993) claims that English speakers sometimes use the terms denoting cardinal directions of the earth (*north, south, east, west*) to describe the same partitions (see section 2.2.1.3 above). According to her research of the dimensional terms only *left* and *right* are used in English for internal regions; however, as noted above in this case there is no way to distinguish a region-based reference system from a speaker-based intrinsic one; even the mention of *side* may implicitly refer to the speaker's, not the scenario's sides. Interestingly, Carroll also claims that English speakers generally prefer intrinsic reference systems that build on the internal features of objects present, while German speakers prefer to partition the perceived scene into internal regions from their own point of view, ignoring the intrinisic features of the objects present. This rather far-reaching claim – involving assumptions about language-specific differences in 'thinking for speaking' as well as inbuilt constraints in the languages with respect to the representation of spatial situations – has not, to my knowledge, been confirmed or challenged in a broader range of discourse situations.

Note, however, that the option of partitioning a certain region within the visual field into subparts may coincide with referring to the internal regions of a relatum such as a room, as described in section 2.2.2.1, notably, if the observer's view originates at the front region of this region or relatum. For example, *hinten im Raum* and *in the back of the room* (based on the room's intrinsic features, such as the position of the entrance) is in effect equivalent to *hinten* based on distance from the observer's viewpoint if the observer is positioned at the entrance. Another case of coincidence occurs when the internal regions of a picture are described, using the outside view, cf. section 2.1. Here, again, the perceived region is divided into subregions based on the observer's viewpoint. In contrast to the regions of the visual field or a certain delimited region inside the visual field, however, pictures (being two-dimensional and often viewed upright, as on walls or on a computer screen) are usually described using the lateral and vertical axes, instead of all three axes. Furthermore, what is described using the vertical axis in pictures often corresponds to the perceived frontal axis in the real world.

In such cases, i.e., if the sides of the relatum are only ascribed through the observer's view direction, then both kinds of reference systems correlate such that English usage does not seem to differ from the German case any more, as data presented in Gorniak & Roy (2004) show. They report participant's spontaneous descriptions of single objects within a scene containing many similar objects presented in 3D on a computer screen. Here, utterances like *the green cone in the back* are used for referring to objects most distant from the observer.

This parallels the usage in German reported by Carroll (1993) in that the scene is partitioned into several areas using dimensional terms. Thus, it seems that the nature of the difference in usage between English in German is less clear than previously assumed.

2.2.2.3 Differentiating internal and external relations: an illustration

In the situation depicted in Figure 5 the adult's *right* (not *left*) arm is patting the child's shoulder, which exemplifies the inside view of the adult. This conception (used for internal relations and descriptions of parts only) coincides with an external intrinsic perspective, as in:

(44) The child is standing to the right of the adult.

where the child's (external) position is described on the basis of the adult's intrinsic features.

Figure 5. Internal vs. external relations

However, an outside observer standing opposite the two people could still say:

(45) From my point of view, the child is standing to the left of the adult.

using a relative reference system. This description, in turn, coincides with taking the picture as an internal reference frame, since pictures are viewed from the outside:

(46) The child is in the left side of the picture.

Finally, the description could also pertain to the observer's visual field:

(47) The child is on the left (hand side).

(48) Das Kind ist links / auf der linken Seite (meines Blickfeldes).
[lit.: The child is left / on the left side (of my visual field).]

Thus, different and (on the surface) contradictory descriptions can be used for the same situation, depending on whether internal or external relations are described, an on whether the objects involved are viewed from the inside or from the outside.

In the present work, in general I do not assume that specific syntactic forms correspond to specific conceptions and reference systems. At least, the empirical evidence for such an assumption is much too sparse for any such conclusions. For instance, adverbs such as *links* in German are potentially ambiguous since (in the present example) they could refer to the visual field as well as to an underspecified relative reference system, etc.

2.2.3 Topological distinctions expressed linguistically

According to Herskovits (1986:173), the application of the dimensional preposition groups (as represented in Figure 3 above) is constrained by their ability to express not only internal versus external relationships, but also a third topological option, namely contact. Landau & Jackendoff (1993:229) describe a very similar categorisation of regions pertaining to spatial expressions other than dimensional terms:

> "[D]istance is digitized into several discrete categories. There appear to be four levels described by English prepositions: (1) location in the region interior to the reference object (*in, inside*); (2) location in the region exterior to the reference object but in contact with it (*on, against*); (3)

> location in the region proximate to the reference object (*near*); and (4) location distant from the reference object (*far* and perhaps *beyond*)."

This distinction is reminiscent of more abstract partitions of space that have been studied extensively in the Artificial Intelligence community, such as the Region Connection Calculus (RCC-8, e.g., Randell et al. 1992) which formalises connections between regions on the basis of eight basic relations. With respect to language, conceptual neighborhood relations (Freksa 1991, 1992a,b) are a useful concept for understanding the linguistic options for dealing with such abstract relationships. Furthermore, Coventry & Garrod (2004a) show in some detail how the RCC approach contributes to understanding the geometric aspects of topological expressions like *in* by providing "a useful means of encapsulating degrees of enclosure across a wide range of types of two- and three-dimensional scenes" (2004a:79).

Following Herskovits, topological relations have consequences for the application of dimensional terms as well. She provides an overview of the dimensional preposition groups with respect to their applicability for the first three of the above categories proposed by Landau & Jackendoff (1993). The distinctions pointed out by Herskovits are well suited to explain, for instance, the difference between (49) and (50):

(49) the child in front of the car

(50) the child in the front of the car

where *in front of* can only be used externally, and *in the front of* only internally. Both terms cannot express contact, in contrast to *on the front of*, which can only be used in cases where contact is involved. Other expressions, such as *on the left hand side of*, are more flexible and can be used internally, externally and expressing contact. Note that Herskovits does not specify whether contact should be internal or external. Intuitively, contact might be a subcase of external relations, in accord with Landau & Jackendoff's second category above. While Herskovits' classification is intuitively appealing I have not yet encountered an empirical validation of it in the literature (but see Talmy 2000:197f. for a similar classification which accounts for Landau & Jackendoff's categories (2)-(4)). Empirical work is specifically asked for with respect to the choice of prepositions in the prepositional phrases that include a dimensional term, since these are the most important distinguishing elements in Herskovits' classification (apart from the employment of an article in *in front of* versus *in the front of*), see Table 4 below.

In cases of group-based reference, it is not always clear whether the underlying reference system should be regarded as internal or external, or (potentially) both, since the target object could either be viewed as being part of the group, or its position could be specified externally, relative to the rest of the group. With the usage of dimensional adjectives as in example (23) above (repeated below as (51) for convenience), this distinction does not seem to yield any difference in interpretation. Example (22) (repeated as (52)) in which the group-based reference is made explicit, in contrast, seems to induce an external interpretation – the target ball is located to the left of the other balls. The distinction is only important when assigning a spatial template in order to determine the extension of the region denoted by the dimensional term (see section 2.3 below).

(51) Kannst du den linken Ball holen?
[lit.: Can you fetch the left ball?; "Can you fetch the ball on the left?"]

(52) The ball is located to the left of the other balls (from my / your / a third point of view).

For German, Klabunde (1999) claims that syntactic variants of the dimensional terms can be differentiated with respect to the external/internal distinction (1999:151):

> "While prepositions like *vor* locate a primary object with respect to a region that is outside the reference object, the corresponding adverb *vorn* locates the primary object in an internal region provided by a reference object. However, a preferred interpretation with an internal relation does not hold for the adverbs *rechts* and *links* referring to horizontal axis-based regions. (…)

> [T]he adverbs *davor, rechts davon* etc. have different meanings based on an external region again".

Klabunde does not offer any proof for his claims either. Table 4 below integrates Herskovits' (1986:174) and Klabunde's (1999:153) proposals, combining the categorisation of regions with the axial distinctions. From this overview, it becomes apparent that expressions for lateral contact only appear to be missing in both languages, as well as terms that flexibly express either internal, external, or contact relations on the frontal axis.

Table 4. Realisations of horizontal relations in English (acc. to Herskovits) and German (acc. to Klabunde)

	lateral	**frontal**
external	to/ by the {left/right} {side} of right/ left of rechts/ links {da}von rechterhand/ linkerhand	before/ behind in front/ back of by the front/ back of back of {da}vor/ hinter
internal	in the left/ right {hand} side of in the side of	in the front/ back of vorne/ hinten
contact		on the front/ back of
internal or external	at the {left/ right} {side} of on the left/ right of links/ rechts	at the front/ back of
internal, external, or contact	on the {left/ right} {side} of	

Ehrich (1985:147) observes that the German adverbs *vorne* and *hinten* lose their intrinsic reading if used together with primary deictic expressions:

> "*Hans sitzt vorne* heißt auf ein Kino bezogen immer, dass Hans an der Leinwand nächst gelegenen Seite sitzt. *Hans sitzt hier (da) vorne* kann hingegen auch bedeuten, dass Hans in einer der hinteren Reihen des Kinos sitzt; Voraussetzung ist dann, dass diese dem Betrachter näher gelegen sind als die vorderen Reihen."
>
> [*Hans sitzt vorne (Hans is sitting in (the) front)* in relation to a cinema always means that Hans is sitting in the region closest to the screen. *Hans sitzt hier (da) vorne (lit.: Hans is sitting here (there) in front),* in contrast, may also mean that Hans is sitting in one of the back rows of the cinema; a pre-requisite for this is that these rows are closer to the observer than the front rows.]

Thus, according to Ehrich expressions like *vorne* are interpreted intrinsically with respect to a relatum like the cinema, and expressions that combine deictic terms with dimensional adverbs are interpreted deictically with respect to the observer. In the present framework which does not use the intrinsic / deictic distinction because it fails to account for a number of relevant distinctions, interpretation is as follows. In both cases internal relations are expressed, but with two different kinds of relata. In the former case, the relatum is the cinema, and in the latter case, it is the observer's visual field (see section 2.2.2.2). Then, an expression like *vorne* can be used in any situation, i.e., without necessitating a situational framework such as that offered by the cinema.

Carroll (1997:142) concludes from experimental work that at least *oben* in German can be used to denote both internal and external relations, as opposed to *above* which is only used for external relations[11]. This observation could be taken to supplement the accounts of Klabunde and Herskovits, who concentrate on the horizontal directions. Additionally, the results of Carroll's work suggest that the frontal expressions in English that can be employed for

[11] She also claims that, on the other hand, *top* can only be used internally; but it is not entirely clear which syntactic form is intended here (possibly as part of *at the top*). *On top of* surely denotes external relations.

internal relationships are not applicable to the partitioning of the visual field as described in section 2.2.2.2 above. Possibly, this kind of usage is restricted to German (and, of course, possibly other languages that are not considered here).

Eschenbach et al. (2001) show how German dimensional adverbs can be combined with spatial prepositional phrases denoting internal regions, rather than with an external relatum:

(53) Die Teller stehen oben im Schrank.
[lit.: The plates are standing above in the cupboard.]

(54) Das Auto steht hinten in der Garage.
[lit.: The car is standing at the back in the garage.]

Their observations basically support Klabunde's suggestions, namely, that German dimensional adverbs are used exclusively for internal relations. The argument of the adverbs may be expressed in a prepositional phrase that itself contains a spatial expression. Alternatively, it can be omitted and be derived contextually.

This possibility leads to a semantic and structural ambiguity: Example (53) could either be interpreted with respect to the cupboard as a bounded internal region, or it could be interpreted with respect to another region that needs to be derived contextually, as shown in (55) where the PP *im Schrank* is not the reference region of the adverb, but rather the PP *in der zweiten Etage* which (in an alternative interpretation) could have been implied in (53).

(55) Die Teller stehen oben in der zweiten Etage im Schrank.
[The plates are standing above on the second floor in the cupboard.]

Eschenbach et al. (2001) point to similar ambiguities with regard to combinations of dimensional terms, which involves additional interpretational problems that will not be pursued here.

Carroll (1993:27) claims that English imposes constraints on the usage of dimensional terms with respect to internal regions of objects, juxtaposing *in front*, which can be used both externally and internally, with *in front of the car* which denotes the external region of the car. She concludes that German is more flexible than English, since by varying *vorne* with *vorne im Auto* the relatum of an internal region can be given explicitly, which according to her is not possible in English. It is an interesting observation that German dimensional adverbs take prepositional phrases denoting internal regions (*im Auto*), while the English counterparts take non-spatial (relational) prepositional phrases (*of the car*). However, Carroll seems to ignore the possibility of using *in the front of the car* (see example (50) above), where the relatum of an internal region is given explicitly. This observation casts some doubt on Carroll's claims with respect to constraints in linguistic spatial representations in German and English, especially because she generally assumes a direct correspondence between syntactic forms and underlying reference systems which is not convincingly motivated. In any case, in order to map her findings onto the present framework, a re-analysis of the data would be necessary because she does not distinguish consistently between perspectives and reference systems (see section 2.2.1.4 above).

2.2.4 Environment-induced reference systems

Up to now, all reference systems could be said to have been used to represent relationships in the real world with its canonical relations: the vertical is defined by gravity, and other directions are derived from it.

Imagine, however, a situation in which real world conditions are less relevant than environment-centered conditions: for example, consider a picture lying on a table, depicting a situation involving a vertical direction (such as people standing upright); as the picture is lying on the horizontal plane the depicted vertical orientation does not correspond to the vertical orientation in the real world. Talking about the picture would surely involve reference

to the vertical with respect to the environment depicted there, not with respect to the real world. A different situation involving human observers could be a dive inside a sunken ship, whose ceiling is oriented at a certain angle with respect to the usual upright direction, etc. Especially underwater, where gravity is not felt as intensively as on solid earth, people might well adopt the directions from the ship's tilted environment, creating a new "vertical". The results by Friederici & Levelt (1990) can be explained in precisely these terms (cf. section 2.2.1.4): in the absence of gravitational cues, a new vertical is created and reference systems employed on that basis (in this case, on the basis of either the head retinal or the visual background). All these situations can be viewed as a kind of *pretense*: the environment is treated *as if* it corresponded to canonical conditions in the real world. Then, reference systems are employed as usual.

2.2.5 Summary of reference systems: Systematic overview

As indicated initially, the literature does not provide a systematic comprehensive overview of possible reference systems and their relationships to perspectives. Figure 6 sums up the findings of the present section. It is restricted to those variations in which spatial dimensional terms are used; thus, absolute reference systems are not included because they rely on different linguistic material. Nevertheless, the network represents conceptual-semantic relations and therefore does not make any claims about linguistic realisations. It is assumed here that, in general, specific linguistic forms do not map directly on specific variants of reference frames. However, some suggestions for linguistic restrictions are represented in Table 4 above.

The network should be read as follows. Since dimensional terms are relational, they always involve two entities to be related, which are here called referent and relatum. First, an axis (lateral, frontal, or vertical) and a topological relation (internal or external) need to be determined. Each of these decisions have further consequences. Internal relations concern either the region defined by the visual field of somebody (usually either the speaker or the addressee), or some entity that either possesses intrinsic features, or is ascribed sides in an *ad hoc* way, or is currently in motion. If the latter is not the case, then the regions can be defined by conceptualising either a view from the inside, so to speak, or from the outside (see section 2.2.2.2 above).

External relations rely, on the one hand, on a reference system (intrinsic or relative), and on the other hand, on perspective. Perspectives can be defined by persons or other animated beings, by locations, and by motion. Furthermore, they can be determined via objects which either have features that resemble perceptual organs, are conceptualised as oriented towards the related entities even if they do not possess intrinsic orientation (as in example 22 above), or by specific object functions that specify directions (as with a car that usually moves in a certain direction, or a pointing arrow).

Intrinsic reference systems can be further specified in that, in the case of *in front*, also the referent can define the front side. In all other cases, the relatum's sides are decisive (see section 2.2.1.1 above). Furthermore, if object functions define an intrinsic relationship then outside or inside view directions can be adopted. This has only consequences with regard to the ascription of left and right, but not with frontal or vertical axes. Since outside and inside views are nevertheless always conceivable no matter which axis is actually employed, this is not viewed as a constraint and therefore not included in the network. Finally, relative reference systems can be either group-based (relying on objects of the same kind as relatum) or landmark-based (relying on an object of a different kind as relatum).

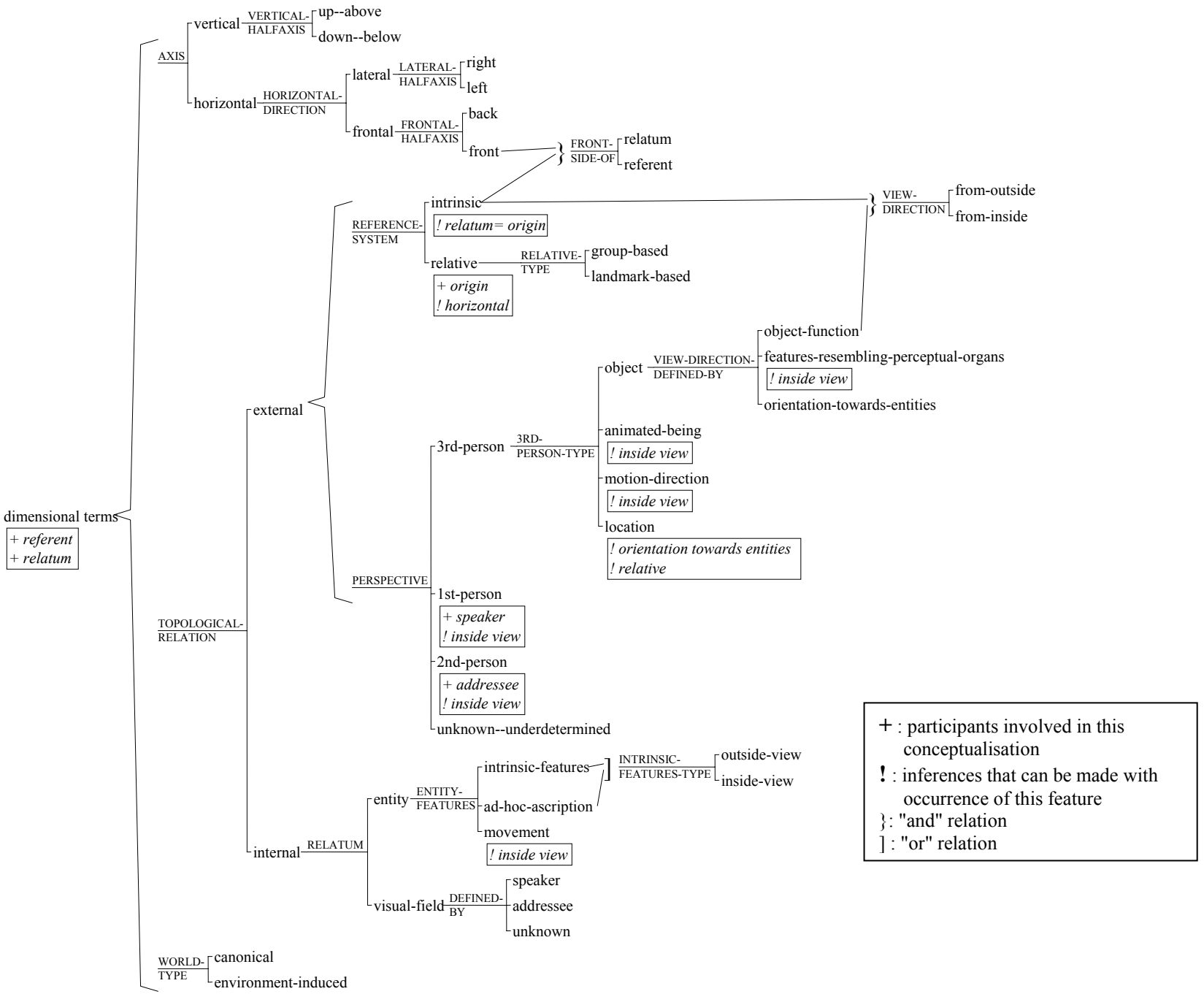


Figure 6. The usage scope of dimensional terms: Options of reference systems, axes, and perspectives

It should be noted that intrinsic object features come into play in two different ways: on the one hand, internal regions can be defined by object features, and on the other hand, a view direction is sometimes derived from them. In the former case, it is the *referent's* intrinsic features that are relevant, and in the latter case, the *origin's*. Furthermore, the view from outside or inside can play a role both in determining internal regions and in defining an intrinsic reference system that builds on the intrinsic features of an object that can be viewed either from the outside or from the inside.

Movement always induces a view from the inside, just as the adoption of a person's perspective (be it speaker, addressee, or a third person) does. This is relevant in the partitioning of an entity's internal parts as well as in the adoption of a perspective induced by direction of motion. This kind of dependency between features is captured via an exclamation mark (**!**) to specify which kind of feature can be automatically inferred whenever a certain feature occurs. Thus, relative reference systems can only occur with the horizontal axes, as explained in section 2.2.1.4 above.

The network presented here builds the basis for an integration in the spatial ontology developed in the project I1-[OntoSpace] (see e.g., Bateman & Farrar 2004).

2.3 Spatial templates

The present chapter started with reflections on axes (section 2.1) followed by a systematic account of underlying reference systems (section 2.2). The generalisations made so far have abstracted from a number of phenomena which will now be examined. The present section addresses the fact that the entities related by dimensional terms do not need to be situated directly on one of the axes with respect to a reference system. Instead, they must be *related* to the axis in some way: In one view, they must be situated within a *region* that is defined by the axis denoted by the dimensional term. In a different view, such as that proposed by Herskovits (1986) (section 1.3), the semantics of the dimensional terms contain a gradedness with regard to applicability, similar to the kind of prototypicality proposed by Rosch (1978). In other words: the closer the object is to the axis, the better the fit to the dimensional terms. The terms are not discrete and mutually exclusive, but applicable to different degrees depending on the distance to the focal axis. The degree of applicability may be influenced further by the overall spatial situation and the discourse task (section 5.1), by functional aspects associated with objects (section 2.8), etc.

These two views are not in conflict. While it may be true that the terms are most straightforwardly applied when the object is situated on the axis itself or very close to it, it may also be true that there is some limit where applicability becomes inacceptable in any kind of situation. The present section therefore does not aim at resolving the two views, but to address gradedness and prototypicality effects as well as delimitations of regions in which dimensional terms are applicable at all. These effects are captured in the notion of **spatial template** as used, for instance, by Carlson-Radvansky et al. (1997). Since a comprehensive and thorough overview of previous insights as well as new experimental results with regard to prototypicality effects of dimensional terms has already been presented by Vorwerg (2001), I will restrict myself to summarizing the most important results. Delimitations of regions are particularly relevant in computational models employed for automatic processing of spatial representations. These will be briefly addressed in section 3 below.

Gradedness effects have been dealt with in the literature in several respects. First of all, psychological research points to the fact that gradedness effects are not restricted to language but can also be observed in non-linguistic tasks (e.g., v. Wolff 2001, Vorwerg 2001:143ff.). Therefore, it is likely that the observed effects are interrelated with degree of cognitive effort. In Franklin et al. (1995) non-linguistic and linguistic tasks are described that were carried out to determine cognitive differences between the various regions in surrounding space. First, the

authors examined the subjective size of each region by telling the subjects, for instance, to "point as far to the front as possible so that you would still consider yourself as pointing to the right". Calculation of the mean values showed that the front region was considered larger than the others, which reflects the special status of the frontal direction compared with the other horizontal directions (see section 2.1 above). In a second, completely non-linguistic experiment, they show that memory for location is more precise in the front than anywhere else; the degree of deviation in accuracy increases with the degree of deviation from the frontal axis. A third experiment then showed that linguistic descriptions are more precise for the front region. Interestingly, the term *front* is treated as a default that can even be omitted; thus, a description like *a little bit to the right* means that the entity's location is very close to the *front* (not *right*) axis. Altogether, these results show the consistent interplay of cognitive and linguistic effects.

They also show something else, namely, that subjects do not simply divide surrounding space into four regions of equal size, as could be expected if all axes were cognitively equal, and all dimensional terms simply exhibited the same gradedness structure. Instead, each region is conceptualised differently, with the frontal region exhibiting privileged status. It seems, however, that these results do not apply to other kinds of tasks that do not involve the human body. Thus, it can be concluded that size of region is a non-stable cognitive variable that varies with spatial task as well as axis.

Crawford et al. (2000) propose a strikingly different interpretation of results that are essentially similar to those of other authors. In their interpretation, the prototypes of linguistic categories are **boundaries** in non-linguistic spatial categorisation. Thus, they offer a different theoretical framework for results that are consistent with both views. However, this alternative interpretation does not seem to have been accepted widely in the literature, though the work is often cited as a contrasting account.

Turning to the area of linguistics, with regard to the application of dimensional terms several insights have been established in the literature (no matter whether they are considered as belonging to the semantics of the terms or concern the area of pragmatics by way of determining application, cf. the discussion in section 1.3). Vorwerg & Rickheit (1998:208), who are proponents of the "prototype" view, summarize as follows:

> "There is increasing evidence that spatial categories, such as projective relations, are *not* discrete, mutually exclusive either-or-categories based on critical features with well defined boundaries. Converging experimental and computational results support the idea that projective relation categories are analog, overlapping, internally structured categories based on prototype comparison and with fuzzy boundaries."

For instance, the expressions *left* and *behind* can be used straightforwardly for 90° and 180° angles, respectively. However, the more the angles between the target object and the relatum depart from these focal axes, the more linguistic modifications are used for specifying the spatial relation. Simple expressions are acceptable and applicable in a certain range; outside this range compounds or modifiers such as *left front* or *a little bit to the left* are more typical (Zimmer et al., 1998). These effects are independent of time pressure and secondary cognitive demands (Zimmer et al. 2001). In the empirical part of her work, Vorwerg (2001) describes in detail the graded typicality structure of spatial expressions as categories on the basis of psycholinguistic experiments in which participants are asked to point out the "best fitting" expression for a given spatial relation, to place an object on the basis of a spatial description, etc.

Freksa (1999) discusses the relationship of a **generic** interpretation of terms like *links vor* ("to the left and in front"), in which case the expression can denote any location within an area between the left and front half-axes, to a **specific** interpretation where prototypicality effects can be noted with respect to single instantiations of the term. He points out that descriptions

can be more or less compatible with spatial relationships, and that the availability of other, perhaps more or less detailed, descriptions can influence the applicability of each term. It can be expected that the best fitting expression out of the available repertory will be chosen. This repertory is influenced by external factors such as the discourse task, the conceptualised abilities of the communication partner, and so on. In identification tasks, the entity will be identified that is closest to the prototypical interpretation of the description used. By this account, Freksa points to central aspects of the relationship between discourse factors (discussed in section 5.1 below) and findings on prototypicality.

From these findings, the following generalisations can be drawn: First, the most straightforward or "best" case of application is position on the axis (e.g., the front axis) itself. The more the referent's position deviates from this axis, the closer it gets to another axis (e.g., right), and so the more likely a term for the second axis (*right*) will be used. Around the middle area between the two axes, most likely both terms (*front right*) will be used in a compound expression. But *front* can still be used close to the right axis, depending on the situation. In other words, the applicability areas of each term with its neighbor overlap. In the present view, the graded structure is not interpreted as part of the abstract semantics: therefore, what overlaps is not the *concepts* of the spatial relations, not the semantics of the neighboring terms, but only the (spatial) situations in which they can be applied. In other words, though *front* and *left* are probably distinctive terms for speakers as part of the vocabulary, just as speakers probably agree that *green* and *blue* are different colours, there are still situations in which either term can be suitably applied. The largest area that can be assumed (and has been proposed) for applicability of a dimensional term is a half-plane (e.g., Herskovits 1986:181f., Retz-Schmidt 1988, Eshuis 2003). Therefore, while the applicability areas of *left* and *front* are not mutually exclusive, those of *front* and *back* as well as *left* versus *right* surely are. But this seems to be the only hard constraint. Thus, as a working definition, at this point we can sum up as follows:

> A dimensional term denotes a spatial relation between two objects. One object serves as relatum, and the other is positioned within a region surrounding a focal axis with respect to the relatum, based on the conceptualisation of a reference system. The size of the region depends on contextual factors but is at all times limited to a half plane. With unmodified dimensional terms the most likely position is on the focal axis itself; with increasing distance from the axis applicability decreases.

As noted above (e.g., Franklin et al. 1995), in some contexts the frontal region seems to be treated as larger than the others because of the cognitive precedence of the frontal axis. Eschenbach (2004) proposes a further distinction with regard to German dimensional terms. While *über* (over, above) and *unter* (under, below) are subject to gradedness effects, *oberhalb* and *unterhalb* are not:

> "The experiments do not cover the prepositions *oberhalb* ('above', 'higher up') or *unterhalb* ('below', 'at a lower level'). However, judgments of native speakers give evidence that *oberhalb* and *unterhalb* are neutral in regard to alignment in the vertical dimension."

To account for this difference, Eschenbach distinguishes between a **strict** and a **lax geometric bearing**, thus including a strict geometric relationship in the semantics of *über, unter* and a relaxed one for *oberhalb, unterhalb.* This proposal, however, does not reflect gradedness in applicability.

Alternatively, it is conceivable that the difference with respect to gradedness requirements is not a binary distinction but itself rather a matter of degree. Thus, the acceptability of spatial deviation from the focal axis may be much more flexible with *oberhalb* and *unterhalb* than it is with the basic prepositions, but basic prepositions – as is well known – can still be used outside the focal area (called **strict geometric bearing** by Eschenbach). Such speculations need empirical validation; intuitions (even collected intuitions by a number of native speakers) are clearly not sufficient here.

Furthermore, Eschenbach observes that, like *oberhalb/unterhalb*,

> "[t]he use of the adverbs *oben* and *unten* and the adjectives *ober* and *unter* does not require alignment of the figure and the contrasting location or ground. Thus, these lexemes are specified with reference to the lax geometric bearing of the reference system."

Here, the same cautions as indicated above apply.

In addition to the effects of distance to the focal axis, there is a further effect of the distance to the relatum (Hayward & Tarr 1995). The acceptability of dimensional terms at some distance from the axis increases with distance between the objects. Thus, the angle between the two related objects seems to be important. On the axis itself, however, distance does not have any effect on applicability. With regard to this latter finding, Landau (2003) reports systematic differences between linguistic applicability and cognitive effects: In language, monomorphemic terms are preferably used along the axes but not outside of the axes, independent of distance; while in memory, there are graded effects of distance both with respect to the axis itself and with respect to distance away from the axis.

These observations pertaining to the features of spatial templates all apply for distance on a small scale, i.e., the experimental results are based on a scenario in which all objects are situated within a region within the visual field. Further aspects of proximity versus distance are discussed in section 2.7 below.

Gapp (1995) addresses a further factor influencing the applicability of dimensional terms, namely, the extension of the relatum. His experiments show (1995:9):

> "The larger the extension of the reference object perpendicular to the canonical direction of the relation, the larger the relation's region of applicability in this perpendicular direction."

Thus, although many accounts in the literature treat the related objects as point-like, their shape is decisive for the applicability of dimensional terms (see also section 2.4 below).

It should be noted that all of these studies deal with **qualitative** descriptions of the spatial relations. It would be interesting to examine the effects of a **quantitative** description on applicability. Consider the following utterances:

(56) The candle is located at 90° to the left of the cup.

(57) The candle is located at 45° to the left of the cup.

(58) The candle is located at 15° to the left of the cup.

Here, it first needs to be decided how the angles are interpreted. Conceivably, the front direction is viewed as 0°. In that case, utterance (56) expresses in metric terms that the candle is situated directly on the left half axis, while utterance (57) denotes a diagonal relationship between cup and candle from the observer's point of view, situated between the front and the left half axes. Utterance (58), in contrast, describes that the candle is situated closer to the front half axis than to the left half axis, in spite of the fact that only the lateral axis is mentioned explicitly. Clearly, such descriptions are unlikely in natural discourse because speakers do not have direct (perceptual) access to accurate measures of this kind. In spatial discourse speakers spontaneously produce qualitative descriptions in which gradedness effects become apparent. But it is unclear so far whether acceptability would increase together with metric measures. Conceivably, example (58) would only be acceptable together with a second dimensional term, indicating the frontal axis explicitly although it can be inferred by the angle information; similarly, acceptability could be increased for (57) in this way.

2.4 Spatial properties of the objects involved: Abstraction, Boundedness, and Extension

Herskovits' (1986) description of the semantic core of dimensional preposition groups starts from idealized conceptualisations of geometric points rather than extended objects. This approach, which has also been followed so far in the present work, corresponds to the insight

that the application of dimensional terms does not pose any specific constraints on the spatial properties of the involved objects, in the way *along* does, which requires an elongated relatum (Talmy 2000, Landau & Jackendoff 1993:226). This fact makes it possible to a certain extent to abstract away from the spatial features of real-world objects. Intrinsic reference systems rely on the relatum's axial structure and/or its axially determined internal regions. However, the presence of intrinsic features is no requirement imposed by dimensional terms because – as long as an origin, an observer's view point, can be determined – it is always possible to employ a relative reference system that is independent of the relatum's properties. This distinction between concrete situations versus abstract geometric structures (imposed frames of reference) is also addressed by Eschenbach (1999:332) who introduces a terminological differentiation between **spatial reference systems** and **spatial frames of reference** which, however, is unique in the literature as far as I can see.

However, the fact that dimensional terms do not require any specific object properties does not entail that such properties do not influence application. Herskovits (1986:184ff.) points out that, in the application of the terms, the extension of the real-world objects as well as their (un)boundedness in any direction does play a role. For instance, if the relatum (reference object) is a square rather than a (conceptualised) point, the focal axes that are relevant for the description of each of the dimensions (on the horizontal plane) are also extended. This could (depending on the underlying reference system) look as depicted in Figure 7. Depending on other factors, application of the dimensional terms is also possible outside the focal regions, mirroring the gradedness of applicability when assuming focal axes with (idealised) points as relata.

In contrast to this view, most approaches tend to treat even extended objects as point-like, using the central axis of each region as focal axis and treating the surrounding area as the field of possible application with gradually decreasing applicability. This approach captures the fact that objects' "center of mass" is a likely focal point serving as a basis for spatial relations. Apart from the center of mass, functionally salient parts of objects can serve the same purpose (Carlson et al. 2003). In both cases, the object may essentially be treated as point-like, with differing criteria used for determining the focal point of the object in each case. Formal models have tried to capture such variation to different degrees (section 3).

Herskovits (1986) also deals with objects that are not as symmetric as that depicted in Figure 7. In general terms, she describes the attribution of spatial regions as follows (1986:185):

> "Enclose the object between parallels to the base axes, so that each such parallel is tangent to the object. The right hand (respectively left hand, front, or back) strip extends toward the object up to the part of the boundary limited by the tangency points first encountered. This procedure requires the object to be bounded so it can be enclosed between those parallels, and thus does not apply to roads, rivers, etc."

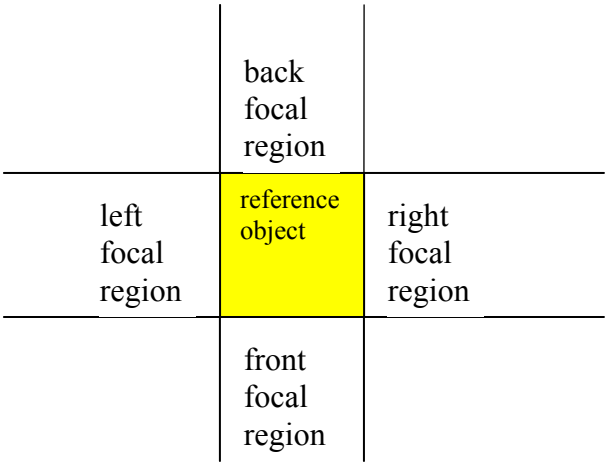


Figure 7. Spatial regions of an extended relatum

Similar ideas can be found in the more recent literature, with different terminology. Carlson et al. (2003:124) refer to the "strip" as "grazing lines" (with respect to the vertical axis):

> "(…) placement of the located object above or below a horizontal line running through the topmost point of the reference object. We refer to this line as the **grazing line**."

Similarly, a closely fitting "bounding box" around an object with curved axes may serve as an approximation for determining spatial regions (e.g., van der Zee & Eshuis 2003).

If, following her definition, regions overlap, Herskovits predicts that conjoined expressions will be used, as in the situation depicted in Figure 8 where *to the right and behind* is a likely description:

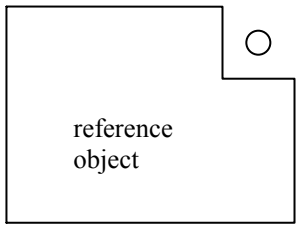


Figure 8. Object configuration with overlapping focal regions

However, it is difficult to imagine in what kind of situation this description would be likely to occur. In most real-world discourse situations further factors, such as functionality of the objects involved (see section 2.8 below), come into play that influence the linguistic choice to a higher degree than would be expected considering Herskovits' suggestion.

With unbounded linear objects, such as roads, reference is possible only with respect to either the frontal or the lateral axis at a time, dependent on the observer's viewpoint (Herskovits 1986:186). Thus, if the observer is standing on the road, looking ahead, a description like *to the right of the road* makes perfect sense, since the road is bounded in the described direction with respect to the observer. However, this is not true in a situation in which the observer is standing beside the road and looking toward it at a right angle, since there is no right region in such a situation due to the perceived unboundedness of the road.

So far, we have considered different spatial properties of the relatum, as opposed to the referent that is still conceptualised as a point. This approach corresponds to a generalisation Talmy (2000:187) proposes, namely:

> "the seeming majority of spatial elements schematize the figure solely as a point or related simple form, in contrast with the treatment given the Ground."

As a counterexample to this general tendency, he shows that the path-related expression *across* depends on rather specific features of the Figure as well. But usually, due to the specific generalizable relations between Figures (referents) and Grounds (relata) (see section 5.2 below), the Ground's features are more relevant for spatial reference.

However, Talmy's observations concern speakers' tendencies for choosing Figures and Grounds in discourse, rather than the attribution of regions once the choice is made. Herskovits (1986) considers some abstract cases in which both objects that are described relative to each other are extended (which, of course, cannot be excluded for real discourse situations). In such a case, the respective spatial regions attributed to them may (partially) overlap. Here, also, conjoined expressions are expected depending on which of the objects is larger and how much overlap there is (see suggestions of suitable descriptions as suggested by Herskovits 1986:187). Furthermore, the order of conjoined expressions depends on the kind of overlap:

> "[I]f A extends partially over the right hand strip, one says *to the right and behind;* if over the back strip, then it is *behind and to the right*."

It may also matter whether the whole extension of the located object is described, as in a description like (59):

(59) The garden is to the right and behind the house.

Another possibility is that only salient parts of the located object are selected for the spatial description, as in example (60) where not the whole river is supposed to be in front of the observer, but only that part which the observer is looking at.

(60) The Thames is in front of me.

Finally, bounded located objects may extend along a portion of an unbounded linear relatum, as in

(61) the land to the right of the road

depending, again, on the observer's position and viewpoint.

Since Herskovits' work, the impact of the involved objects' extensions has been addressed empirically by several authors. Vorwerg (2001), for example, reports several experiments involving various extended (rectangle-shaped) relata together with a small ring as referent. Results show that orientation and shape of the relatum relative to the observer's view direction systematically influence the application of dimensional terms. Vorwerg examines these effects predominantly with respect to the expressions' graded typicality structure (see section 2.3 above). In Vorwerg (2003), she shows specifically how the relation of the relatum's longitudinal axis to the view direction influences choice of dimensional terms ("tilt effect"), thereby proving the impact of perceptual aspects (Vorwerg 2003:328):

> "As the tilt effect can be interpreted as a perceptual effect, its effectiveness in different language tasks, such as the interpretation of a spatial expression or the use, applicability rating or verification of a spatial term for a given direction, can be regarded as strong evidence for a perceptual foundation of linguistic categorization of space and the use of perceptually based reference directions as prototypes or comparison values for linguistic categorization."

As van der Zee & Eshuis (2003) show, object shapes also determine the assignment of intrinsic reference axes (see section 2.1 above). Furthermore, Coventry & Garrod (2004a:55) correctly point out that ignoring most of the object features (apart from a schematized representation that is required by each spatial term), as Talmy (1983), Landau & Jackendoff (1993) and others propose to do, leads to a failure to appreciate the influence of object functions on the application of spatial terms (cf. section 2.8 below).

2.5 Conceptualisations and granularity

In some cases, conceptualisations of object extensions differ from the actual shape of objects in the real world. Thus, Landau & Jackendoff (1993:222) distinguish between **surface-type** and **volume-type** objects. Surface-type objects are conceived of as two-dimensional, while volume-type objects are treated as three-dimensional. If a surface-type object, such as a (gramophone) record, is described by an attribute like *big*, the description applies only to the horizontal extensions, but not to the vertical dimension, since records are conceived of as two-dimensional. In contrast, a "big house" is usually conceived of as "big" in all three dimensions.

Whether we conceive of an object as one-, two-, or three-dimensional may be a matter of purpose. As Hobbs (1995:820) points out, with reference to a road:

> "When we are planning a trip, we view it as a line. When we are driving on it, we have to worry about our placement to the right or left, so we think of it as a surface. When we hit a pothole, it becomes a volume for us."

Grabowski (1998) notes that sometimes different conceptualisations of the dimensionality of the same objects can be used in the same discourse, as evidenced by the spatial expressions used. Thus, an expression like *das Dorf liegt an der Hauptstraße* (*the village is at the main road*) treats the village as point-like and the street as linear. But in a sentence like

(62) An der Straße, auf der ein plattgefahrener Igel liegt, befindet sich das Dorf, in dem die Kirche steht.

[At the street, on which a roadkilled hedgehog is lying, there is the village in which the church is.]

the street appears both as linear and as surface-like (planar), and the village as point-like and as container-like, since a church is positioned within it. Thus, assigning abstract spatial relations to linguistic representations can be a complex task, which is solved by flexibly producing suitable conceptualisations for different purposes. However, a sentence like Grabowski's constructed example (62), which densely combines a number of different conceptualisations, seems to be unlikely to occur in natural discourse. How different conceptualisations of spatial settings are combined spontaneously by speakers needs to be addressed empirically.

Such variability in conceptualisation can also play a role in the application of dimensional terms. Since the applicability areas vary with the extension of the object, different conceptualisations of actual configurations can lead to different interpretations of the same situation. For instance, imagine a tree that could be conceptualised as three-dimensional, as two-dimensional (ignoring the vertical dimension, as is often the case in the application of the horizontal expressions), or as one-dimensional (ignoring the extension on the horizontal plane because of the salience of the trunk). Now, an utterance like

(63) The bird is in front of the tree.

could receive any of three different interpretations, depending on the conceptualised dimension: In a three-dimensional conception, the bird could fly anywhere between the tree's extension and the observer. In a two-dimensional conception, the bird sits on the ground between the tree's horizontal extension (including branches and leaves) and the observer. In a one-dimensional conception, the bird is situated on the ground between the trunk and the observer. Alternatively, it is possible also to ignore the frontal axis and refer to height and/or the lateral axis only, again reflecting a two-dimensional rather than three-dimensional conceptualisation. As discussed in section 2.1, it seems to be common to interpret dimensional terms two-dimensionally, either ignoring the vertical or the frontal axis.

These phenomena are closely interconnected with the concept of **granularity**, as Bateman & Farrar (2004:94) point out. They use a different, suggestive example:

> "From a distance, the beach looks like a smooth surface, but when lying on it, it is distinctly grainy; the surface of a table usually looks very smooth, but viewed in terms of its molecules, the boundary is very much harder to describe as anything resembling smooth."

Granularity concerns choice of aspects that are relevant or significant in the current discourse. This could be a matter of size, as illustrated by the following two examples:

(64) Harburg is close to Hamburg. (underlying granularity level: German locations)

(65) München is close to Hamburg. (granularity level: cities of the world)

In such cases, the level of specificity is important for interpretation of the spatial descriptions. Likewise, a description like

(66) Harburg is ten kilometers from Hamburg.

does not suggest that the distance between Hamburg and Harburg is supposed to be exactly 10.000,00 m (plus zero meters, centimeters, millimeters, etc.), but rather, that it is conceptualised as something close to that estimate. Interpretation is usually generous in the sense that it allows for deviations of up to several kilometers without rendering the utterance untrue. However, if a more specific level of description is asked for, an utterance like (66) could well be considered as uninformative or even false. Thus, granularity levels apply even in situations involving quantitative instead of qualitative descriptions. Metric measures do not need to be precise, depending on the level of granularity involved.

With qualitative descriptions involving dimensional terms, levels of specificity may also play a role. Thus, there may be situations in which a rough description like

(67) The ball is to your right.

is sufficient, for example, if the interlocutor has full visual access and the ball is not far away. But in a different situation, for instance, if the interlocutor is blindfolded and the ball is at some distance, utterance (67) is not informative, since more precision is needed. Even if the ball is located directly on the right half-axis, a precise (though qualitative) description like

(68) The ball is exactly to your right.

would be much more suitable. Thus, what is relevant in a situation depends on functions and purposes, and this will influence the interpretation of any spatial description.

One aspect which is certain to influence the level of specificity as well as the kind of spatial description chosen is the presence of other objects nearby that represent competing candidates for reference. Pribbenow (1992:160), for instance, states that for a spatial description to hold it is necessary that there is no more direct relationship to another object in a given situation. Such aspects will be dealt with in more detail in section 5.1 since they are intricately connected with discourse-related factors such as the question under discussion or "quaestio", i.e., the given discourse task.

As example (63) above and its diverse interpretations show, however, relevant aspects chosen for spatial descriptions need not always concern specificity, but can also relate to other features such as dimensionality. This fact has for some time been recognised in work on spatial ontologies, as described, for example, in Bateman & Farrar (2004:48ff.). Fonseca et al. (2002) formulate as follows:

> "Some authors consider *granularity* in a spatial database to be the same as resolution, thus implying that granularity is related to the level of distinction between elements of a phenomenon that is represented by the dataset. (…) Resolution refers to the amount of detail in a representation, while granularity refers to the cognitive aspects involved in selection of features."

As shown in this section, such "cognitive aspects" can concern the level of specificity as well as the choice of relevant dimension. But in theory any kind of feature might turn out to be relevant in a situation. Further aspects influencing relevance and salience, and thus choice in terms of granularity, are interconnected with the functionality of the objects involved, which is addressed in more detail in section 2.8.

2.6 Accessibility

One of the major features that distinguishes (spatial) objects from (temporal) events is the fact that objects are directly accessible in a situation. For example, time cannot be explored actively. This allows for much more varied interaction with the spatial domain in contrast to the temporal. Spatial relationships can be expressed linguistically by implicitly or explicitly relying on the present position of the speaker. The most obvious example is the usage of expressions such as *here* which are described as **deictic** by Levinson (2003). In this view, deictic expressions do not employ a frame of reference, and the spatial relation is described "in terms of radial categories ('here' vs. 'there'), or in combination with a pointing gesture ('there' with a point)" (Levinson 2003:65). Other authors also describe the dimensional terms as deictic. While, as motivated in section 2.2 above, I do not adopt the deictic / intrinsic distinction that is often encountered in the literature, in the present section I address the underlying reason why dimensional terms are often classified as deictic, namely, their situation dependency and accessibility.

This feature of dimensional terms is directly reflected in Grabowski's definitions (cf. p. 14) in which perceptual factors play a decisive role. On the front half-axis, the object is located directly in the visual field of the observer; on the lateral axis, the observer only needs to turn

their head to perceive the located object. On the back half-axis the located object is occluded either because it is behind the observer (in an intrinsic reference system) or because it is behind the relatum (in a relative reference system). Thus, two kinds of *causal* relationships occur between perceptual accessibility and application of dimensional terms (Grabowski 1998):

> "Für Innen-Origines ist das Zielobjekt sichtbar, *weil* es sich in Sichtlinie der anthropomorphen Origo befindet, oder es ist nicht sichtbar, weil es sich 'im Rücken' der Origo befindet. Für Außen-Origines ist das Zielobjekt sichtbar, weil das Bezugsobjekt dies zulässt (d.h. es nicht verdeckt), oder es ist nicht sichtbar, weil das Bezugsobjekt es verdeckt. Für die perzeptuelle Zugänglichkeit des Zielobjekts bestehen also zwei alternative kausale Attributionsmuster, deren Anwendung von der relativen Position zwischen Origo und Bezugsobjekt abhängt und die damit eine Frage der kognitiv-konzeptuellen Auffassung des beschriebenen Weltausschnitts sind."
>
> [For origins in intrinsic reference systems[12] the referent is visible *because* it is situated in the line of sight of the anthropomorphic origin, or it is not visible because it is situated 'in back of' the origin. For origins in relative reference systems the referent is visible because the relatum allows this (i.e., it does not occlude it), or it is not visible because the relatum occludes it. Thus, with respect to the perceptual accessibility of the referent, there are two alternative causal patterns of attributions, the application of which depends on the relative positions of origin and relatum, and which therefore concern the cognitive-conceptual notion of the described section of the world.]

Note that a reversed relation of occlusion also comes into play on the front half-axis in relative (but not intrinsic) reference systems, where a referent *in front of* another object occludes this relatum from the observer's point of view.

For the frontal axis, (lack of) accessibility was also described as the functional element by Tyler and Evans (2003) as described above. On the grounds of empirical evidence, however, Landau & Jackendoff (1993:231) ascribe only a secondary role to this factor:

> "Visibility and occlusion may well play a role in children's learning of spatial prepositions. Johnston (1984) found that 3-year-olds understood *behind* to mean 'occluded from sight'. For example, they agreed that X was behind Y only when Y was an object large enough to occlude X. In contrast, adults agree that X is behind Y whatever their relative sizes, as long as the figure object is aligned properly with the reference object's front-back axis."

While these results do suggest that children overestimate the significance of the occlusion effect, they do not disprove the hypothesis that the semantics of *behind* at least for some speakers includes an element of *at least partial* occlusion. But note that the effect interacts with the underlying reference system, at least in German: Occlusion effects do not come into play, or at least play a lesser role, with internal regions defined by an observer, such as when the observer's visual field is partitioned into regions (cf. section 2.2.2.2 above). In that case, *hinten* describes a region that is further away from the observer but usually still visually accessible, though maybe with increased effort.

Thus far, we have addressed accessibility in the real world, in a situation in which the objects talked about are present, and the observer is really looking from the origin's position. Of course, nothing of this is a precondition for the employment of spatial expressions, since it is always possible to mentally adopt another person's perspective or the perspective of a non-human entity (see section 2.1); in fact, the origin of a reference system can be any real or imagined position that the speaker may choose for reference.

Furthermore, the whole scenario may not be present but may be represented, for example, graphically, linguistically, or by other means. If spatial relationships are established linguistically with respect to an imagined or imaginable world, the discourse model must provide sufficient information for the listener to be able to identify the relative positions of the

[12] For ease of comprehension in the translation the terminology of the present work is adopted, rather than literally translating Grabowski's terminology.

objects talked about. For the establishment of this information, general discourse-related phenomena apply, such as anaphora resolution, Given/New structure, etc. In addition to these, several aspects pertaining specifically to the linguistic establishment of spatial relationships have been investigated, for example, by Talmy who points to different properties of Figures vs. Grounds (see section 5.2). Klabunde (1999) notes that, in the application of dimensional terms, relata are usually already identifiable for the listener, and therefore generally realised by a definite noun phrase in German (1999:152):

> "Localizations like *hinter einem Haus steht ein Brunnen* (there is a fountain behind a house) with a house as unidentifiable reference object are principally possible, but they do not occur in coherent descriptions. The reason is that the reference objects function as 'anchors' for the single localizations at what has been said before."

Of course, contexts can be imagined in which a description such as that given by Klabunde would make perfect sense; however, he is certainly right in pointing to a general tendency here. This tendency reflects a general fact about natural language, namely, that sentences usually start from some kind of given information and proceed to new meaning that is to be conveyed by the sentence. This difference is often directly reflected by the usage of definite or indefinite noun phrases.

Thus, accessibility pertains not only to objects in the real world but also to objects within the discourse model. With dimensional adverbs used with relative and intrinsic reference systems, relata are often implicit, which leads to the expectation that specifically here, "only identifiable and salient discourse referents should be used as reference object" (Klabunde 1999:150). Thus, omitted relata are expected to be focused in the discourse model. This observation immediately raises the question of what speakers actually do in natural discourse, taking into account discourse task and context. These issues also need to be addressed empirically.

In German, there is a lexicalised option for omitting previously mentioned or otherwise identifiable relata that is not lexically encoded in English, namely, the attachment of the deictic morpheme *da* to either *hinter* or *vor* (yielding *dahinter, davor*). On the lateral axis, a more complex expression is necessary, e.g., *rechts davon.* Klabunde (1999:151) notes that *hinter* and *dahinter*

> "express the same spatial relation. The difference is grounded in the accessibility state of the discourse referent for the reference object and the influence of the meaning of the locative anaphor *da.*"

While these observations pertain to accessibility within a discourse situation, accessibility in purely spatial terms also applies to a discourse that is remote from the actual scene: even if objects are not directly present they should be imagined to be co-present in the scene described linguistically. Thus, the same restrictions as mentioned above with respect to real-world scenarios apply, regardless of whether the objects are actually present or not. The main difference is that, in the latter situation, the language must be more explicit in order to make perceptually non-accessible information linguistically accessible.

2.7 Proximity and distance

Intuitively, in order for a dimensional term to be applicable, the related objects need to be situated within a reasonable distance. In section 2.3 above, it was noted that distance within the visual field may have an influence on the applicability of monomorphemic terms, i.e., with increasing distance greater deviations from the focal axes are acceptable. However, there are some limits to the distance where dimensional terms are applicable at all. The two objects must be conceived of as spatially related, i.e., **proximal** to each other to some degree. Similarly, if objects are used as landmarks in a route description their **influence area** must intersect with the route itself (Gryl et al. 2002). In the literature, this idea is sometimes associated with a concept introduced by Miller & Johnson-Laird (1976) with regard to spatial

relations in general, namely, the notion of a **region of interaction**. This is described as follows (1976:59):

> "In order to take account of spatial relations, the perceptual process must not only register place, but relations between places, which entails perception of a spatial region containing the place of the thing. The region of a thing can be thought of as a rather indeterminate penumbra surrounding it. The advantage of region over place as a perceptual predicate is that regions can overlap even though things cannot. Thus, two things whose regions overlap can be seen in spatial relation to each other. We will say that object *x* is in the region of object *y* when *x* is spatially close enough to *y* to have the sort of interactions with it that normally occur between *x*'s and *y*'s. This definition of region is deliberately vague, because the perceptual attributes of a region are correspondingly vague."

For instance, an expression like *here* can be used to refer to the **personal space** around oneself (although it may also have other, more extended interpretations, as exemplified by *here on Earth*). The conceptualisation of a region of interaction around an entity depends on experience, thus (Miller & Johnson-Laird 1976:59):

> "Once an object has become familiar, its characteristic region of interaction seems to be appreciated with the same immediacy as many simpler perceptual attributes."

This observation is in line with more recent findings on objects' affordances, e.g., by Tucker & Ellis (1998) who find that "seen objects automatically potentiate components of the actions they afford". This line of research goes back to Gibson's theory of perception (Gibson 1979), who claimed that perception is intertwined with potential actions on the environment. Thus, forms and functions of objects are conceptually closely related (a fact which is also exploited in research on automatic object recognition, e.g., Wünstel & Moratz 2004).

Pribbenow (1992:146) points out that the region of interaction – the region within which one object can be perceived as potentially influencing another – is dependent on the features of the objects involved:

> "Der Einfluss des Referenzobjektes ergibt sich direkt aus seiner Funktionalität als Landmarke. Generell gilt dafür, je salienter ein Objekt ist, desto größer der Bereich, in dem eine andere Entität lokalisiert werden kann, wobei sich die Salienz eines Objektes aus unterschiedlichen Faktoren zusammensetzt. In erster Linie interessant sind in diesem Zusammenhang visuelle Eigenschaften des RO, da die LE bei konkreten Suchprozessen primär durch visuelle Prozesse mit dem RO in Beziehung gebracht wird. Dabei ist für die visuelle Salienz neben der Größe die Auffälligkeit oder Besonderheit eines Objektes wichtig, genauso aber auch seine generelle Sichtbarkeit. Ergänzt wird der visuelle Aspekt der Salienz durch subjektive Bewertungen wie den Bekanntheitsgrad der Entität und ihre Relevanz für Hörer und/oder Sprecher der Äußerung. Die letzteren Größen sind wichtig für die Auffindbarkeit eines Objektes und schaffen damit erst die Voraussetzung, dass eine Entität sinnvoll als Landmarke dienen kann. (…) Wie beim Referenzobjekt gilt auch für die zu lokalisierende Entität die Beziehung: je größer das Objekt, desto größer das Suchgebiet. In diesem Falle wäre diese Beziehung treffender formuliert durch: je kleiner die LE, desto näher sollten RO und LE zusammenliegen. Insgesamt gesehen ist es die Kombination der Eigenschaften *beider* Objekte, die die konkrete Ausdehnung des Lokalisierungsgebietes festlegt".
>
> [The influence of the reference object follows directly from its functionality as a landmark. In general terms, the more salient an object is, the larger is the area in which another entity can be localised. Here, the salience of an object is composed of various factors. Of interest in this context are, first of all, visual features of the RO, because in concrete search processes the LE is related to the RO primarily by visual processes. For visual salience, besides size, distinctiveness and noteworthiness of an object are important, as well as its general visibility. The visual aspect of salience is supplemented by subjective judgements such as the familiarity of the entity, and its relevance for hearer and/or speaker of the utterance. The latter factors are important for the identifiability of an object and thereby constitute the prerequisites needed for the entity to be suitable as a landmark. (…) For the entity to be localised, the same relation holds as for the reference object: the larger the object, the larger the search area. In this case, the relation might

be more suitably described by: the smaller the LE, the closer RO and LE should be to each other. Altogether, it is the combination of the features of *both* objects that determines the concrete extension of the localisation area.]

With topological expressions, regions of interaction are undoubtedly relevant. For instance, in utterances like

(69) The house is located at a river.

the interpretation of *at* is certainly dependent on the interactants' concepts of possible interactions between houses and rivers. Now, the question arises whether or to what degree this notion is applicable for dimensional terms. With the terms denoting the earth's cardinal directions, there seems, in principle, to be no limit with regard to distance, as the following example shows (which may be used in a context where one wishes to point out this striking coincidence):

(70) We will spend our holidays at a hotel that is situated exactly south of our home.

However, exchanging *south of* for a dimensional expression does not work, even ignoring the fact that the spatial relations do not necessarily coincide:

(71) ? We will spend our holidays at a hotel that is situated exactly behind our home.

Pribbenow (1992:160) proposes the following general guideline to accounting for the intuition that two objects need to be sufficiently close to each other in order to be related (while pointing out elsewhere that this does not apply for compass directions):

"Das Distanzkonzept der Nähe bezeichnet (…) den Bereich um das Referenzobjekt, in dem das RO als Bezugspunkt für die Lageangabe der LE [der zu lokalisierenden Entität] verwendet werden kann. Es muss also ein Zusammenhang zwischen den beiden Objekten oder, wie Miller/Johnson-Laird (1976) es in ihrer 'region of interaction' andeuten, eine Interaktion zwischen den beiden Entitäten bestehen. Der visuelle Aspekt dieses Zusammenhangs beinhaltet, dass die LE normalerweise vom Referenzobjekt aus sichtbar ist. Die beiden Objekte müssen untereinander zugänglich sein, d.h. es darf keine Unterbrechung des räumlichen Zusammenhanges vorliegen. Die zu lokalisierende Entität darf nicht direkter mit einem anderen Objekt in Beziehung zu setzen sein, d.h. sie darf nicht zu weit vom RO entfernt sein, usw. Der Zusammenhang bzw. die Interaktion zur Lokalisierung hat somit vielfältige Aspekte, die im allgemeinen vom räumlichen Kontext, das ist in diesem Falle die Umgebung des RO, mitbeeinflusst werden. Damit ist die Erstellung eines zutreffenden Lokalisierungsgebietes nur bei Kenntnis der beteiligten Objekte und des räumlichen Kontextes, m.a.W. nur der beschriebenen realen Situation, möglich."

[The distance concept of proximity denotes (…) the area around the reference object in which the RO can be used as a reference point for the localisation of the LE. Thus, there must be a relationship between the two objects or, as Miller & Johnson-Laird (1976) allude in their 'region of interaction', an interaction between the two entities. The visual aspect of this relationship entails that the LE is usually visible from the reference object. The two objects must be mutually accessible, i.e., there should not be an interruption of the spatial relation. The entity to be localised should not be more directly related to a different object, i.e., it should not be too distant from the RO, etc. Thus, the relationship or interaction of the localisation has various aspects, which are usually further influenced by the spatial context, which in this case is the proximate region of the RO. Therefore, the identification of a suitable localisation area is only possible through knowledge about the participating objects and the spatial context, in other words, the described real situation.]

It could be added that discourse-related aspects such as the current task may also be decisive in determining the suitability of relating two objects linguistically, as described in section 5.1 below.

Pribbenow is mostly concerned with topological and distance-related spatial terms, such as *bei* (at), spelling out Miller & Johnson-Laird's proposal concerning the region of interaction that allows for the usage of such terms. She notes (1992:81) that dimensional terms are not described in this way by Miller and Johnson-Laird. It is conceivable that the (original)

concept of region of interaction is itself too restricted for application with dimensional terms, since the notion was developed for narrower areas between entities. However, the criteria proposed by Pribbenow are intuitively appealing and are therefore taken up here, in order to examine in how far they could be hypothesized as a framework for the application of dimensional terms. The most important aspects to note at this point are the proposed constraints with regard to

- visual *access* between the two objects,
- noninterruption of the spatial region between the objects (i.e., no other objects or obstacles should be in the way), referred to elsewhere as *nextness*, and
- *primacy* of the relation between these two objects compared to any other objects present which could serve as relatum.

Formulated in this way, these are rather strong claims that can (and should) be tested empirically. I do not know of any research so far addressing these issues directly, probably due to the fact that empirical studies dealing with dimensional terms usually employ scenarios in which these conditions are met to begin with, for all objects present in the scenario.

In situations where the relatum is not a single object but several objects of the same kind as the referent (i.e., group-based reference), all of the involved objects may need to meet the criteria with respect to each other in order to be conceptualised as suitable for spatial reference. Possibly, the *primacy* condition would conflict with the fact that the referent is related to more than one object, yielding a different situation in which the objects need to be conceptualised as a group, and therefore should be sufficiently close to each other. It would be an interesting question for empirical psycholinguistic research to find out whether the objects need to be close to each other (how close?) to allow for group-based reference, or whether qualitative-functional requirements such as Pribbenow's constraints would indeed be sufficient. Possibly, a formal definition such as that proposed by Thorisson (1994) could be useful here; he defines a proximity score for the distance of each object in the domain to a particular object. The formula is used, for example, by van der Sluis & Krahmer (2000) to determine whether objects are situated within the focus space of another object.

It is important to note that the proposed constraints only apply for the relationship between the objects, not the observer's viewpoint (except in the case of intrinsic reference systems where both coincide). Thus, it is entirely possible to say

(72) The ball is in front of the table.

in a situation in which the ball is occluded from sight of the observer, other objects are located in between the observer's position and the two objects, etc. The only requirement is that the two related objects are sufficiently close in terms of the above claims. (Recall that *behind* often even entails that the observer does *not* have visual access to the referent!)

Another point to note is that the constraints may sometimes be cancelled by metric distances. Thus, an utterance like

(73) The box is two meters behind the barrel.

does not require that no other objects are placed in between box and barrel, etc. Indeed, proximity is not required at all in that case. Thus, the above example (71) might become more acceptable (though still unlikely) by introducing metric measures:

(74) We will spend our holidays at a village that is situated exactly two hundred kilometers behind our home.

Herskovits (1986:84f.) makes a similar point:

> "If I say *The fountain is behind the city hall* you will assume that the fountain is close to the city hall, in fact that it is *next to* the city hall, meaning that no other salient architectural object is between the two objects. This 'nextness condition' is not part of the meaning of *behind*, because in *The treasure is buried 600 feet in a straight line behind you* there is no proximity. Still, we

> find it most useful to relate objects by means of *behind* when they are close together for several reasons: the direction defined by *behind* is otherwise difficult to evaluate, and as it is generally more rational to choose a reference object close to the object one wishes to locate, most typically objects will be close together. As a consequence, 'nextness' is the unmarked case, the default, the inference one should draw, lacking evidence to the contrary."

Likewise, specific discoursal and situational requirements might have the same effect of cancelling the nextness condition, such as, for example, if the position on one of the axes is specifically important (because they are perceptually related), as in:

(75) Could you please place this picture on the wall on a straight line to the right of the upper doorframe at the other end of the room?

Accordingly, it seems to be a good solution to treat the proposed constraints as default conditions which are cancellable by contextual factors, in the same way as many features of dimensional terms can be affected by the situation. One contextual factor that reinforces (rather than cancels) the nextness condition is if a background grid (such as a network of streets or a chessboard) is available (Herskovits 1986:82):

> "If there is in the context a grid-like pattern and if that grid-like pattern is relevant to the speaker's point, then that grid-like pattern will define the tolerance in the use of the projective prepositions."

In general, however, proximity as defined in Pribbenow's terms does not require that the objects are situated particularly close to each other, as long as the criteria are met. Thus, the notorious problem of context dependency in assessing a suitable measure of closeness is reduced considerably, since proximity is not defined in terms of distance but in terms of functional aspects of the situation. Furthermore, Herskovits (1986:75) points out that appearance can be more salient than geometric facts, as when spatial relationships are projected on the plane of view:

(76) The Morning Star is to the left of the church.

Here, the requirements of proximity are met simply because the star (perceptually) *seems to be* situated in a suitable position with respect to the relatum (the church).

I would like to suggest that terms denoting the vertical direction behave differently, which is supported by the fact that the vertical direction is usually based on the absolute system defined by gravity, seldom intrinsically, and (apparently) never by relative reference systems (see section 2.2.1.4 above). Recall, however, that Tyler & Evans (2003) propose a difference between *over / under* and *above / below* with respect to proximity as part of their semantics (see Table 2 above); furthermore, research by Garrod and Coventry points to further influences of functional aspects (cf. section 2.8 below). Altogether, it seems that proximity is more relevant to *over* than to *above*, sometimes involving contact in the case of *over* but not *above*. This idea is exemplified by the associations evoked by the following two sentences (Tyler & Evans 2003:112):

(77) Nora twirled over the polished floor.

(78) Nora twirled above the polished floor.

Here, it seems that Nora must be dancing in the air in the second case. However, according to Tyler & Evans, there is also a sense of *above* which requires the referent to be situated next to the relatum on the vertical axis, as in (p.120):

(79) His office is on the floor above mine.

In light of Herskovits' observation (cited above) with regard to the impact of grid-like patterns, it seems that this effect is due to the features of the relatum (the floor) rather than highlighting a specific sense of *above*. The example shows, however, that in spite of the distality associated with *above,* this term is also sensitive to the influence of grid-like patterns, similar to the other dimensional terms.

The observations concerning differences between *over* and *above* raise the question whether all horizontal expressions show the same requirements. According to some findings in the literature, this is not the case; different authors have pointed to different, more or less subtle differences in applicability of the terms with regard to proximity. For example, Talmy (2000:198) states that **immediate adjacency** is a prerequisite for the usage of some specific preposition groups of dimensional terms, namely, *in front of, in back of / behind, on one side of / beside, on the right/left of*. He claims:

> "The fact that these expressions cannot be used to localize Figures at a greater distance shows that they indicated relative adjacency to the Reference Object. For example, a bike directly lined up with the front of a church but three blocks away cannot be said to be 'in front of' the church."

In contrast to this, a few other expressions can be used to express some amount of distance:

> "(…) the adjacency condition is removed. The Figure is localized in a particular quadrant by reference to some Reference Object part, but it is at any remove."

Talmy suggests that expressions containing *to* rather than *on* belong in this group, as in:

(80) The bike is to the right of the church.

which can, according to Talmy (2000:199) be interpreted to mean "anywhere from three feet to three blocks." This difference between *to* and *on* is well represented in the literature (see also Herskovits 1986:188f.). Talmy adds:

> "*Rearward of* might work for the back direction, as in *The bike is rearward of the church*, but *forward of* will certainly not do for the front direction. In general, conveying these concepts requires lengthy expressions, and then ones that are not neutral to distance but in fact indicate nonadjacency, as in *The bike is a ways off from the front of the church.*"

Another suggestion with respect to different requirements in proximity comes from Landau & Jackendoff (1993:230), who claim that *in back of* and *behind* differ in distance:

> "A tree may be right behind (proximal), way behind (distal), or right in back of a house, but 'The tree is way in back of the house' sounds odd or colloquial (to us, anyway).The standard use of *in back of* seems to be restricted to proximal distance (and possibly contact) whereas *behind* and the colloquial *in back of* are unrestricted."

Herskovits (1986:189) suggests that the use of *at the left / right* imply even closer proximity than *on the left / right*:

> "These imply the closest proximity consistent with a reasonable world: *the bottle is at your left* suggests that the bottle is immediately accessible to hand, while it could be further away with *the bottle is on your left.*"

It can be concluded from this section that, although there seems to be some amount of uncertainty in the literature with respect to the details, there is a general agreement that most dimensional terms require some amount of proximity for applicability. This proximity may be induced by perception or conceptualisation rather than "objective" environmental factors. As described above, Pribbenow (1992) suggests qualitative-functional criteria for the assessment of proximity, which are taken up in the present work as working hypotheses for default conditions cancellable by contextual factors, rather than strong constraints.

2.8 Functional features of objects

As discussed in section 1.3 above, it has been noted by several authors (e.g., Tyler & Evans 2003) that the spatial relationships expressed by dimensional terms are often associated with additional functional features, corresponding to the situations in which they are most often applied. Furthermore, closely related to such findings, proximity has been defined on the basis of functional aspects (section 2.7). The present section deals with a different aspect of **functional geometry** (Garrod et al. 1999): namely, the impact of functional relationships between the *objects* involved rather than the spatial relation itself. While functional aspects of the spatial relationships, such as SUPPORT, SEQUENCE, or ACCESSIBILITY, can be (and have

been) represented in a fairly abstract way, objects' functional features are peculiar to the objects involved. Nevertheless, they may have a strong influence on the application of spatial terms. As Coventry & Garrod (2004a:55) put it: "*what* objects are fundamentally influences how one talks about *where* they are located". Or in the words of Coventry & Mather (2002:182):

> "If one comes across *the x is over the y* then what this means will be determined by object knowledge. If the figure and ground are a *tablecloth* and a *table*, then information about the functions of these objects is available from lexical entries of the nouns accessed, and the spatial relationship between these objects can then be established. If sense selection requires this to be done anyway, then there is no advantage to lexically representing all the senses given the problems that this entails".

This section deals with this kind of influence of object knowledge, guided by the question to what degree the influence of objects' functions can be generalized in order to make predictions about the applicability of dimensional terms. Of course, often both kinds of functional relationships interact to a high degree: a typical situation of employing an expression like *in* (which is generally agreed to involve an associated functional feature of CONTAINMENT) is one involving a *container*, such as a bowl. The applicability of *in* then interacts with the degree to which the functionality of the object as a container is fulfilled (Garrod et al. 1999). That is, in this case applicability depends on a good match between the associated feature of the spatial term, and the functional features of the object(s) involved. As Coventry & Garrod (2004a:61) point out, this may depend on how the object is conceptualised in the first place:

> "Objects are often associated with particular functions, and retrieval of this functional information from memory may promote the application of different routines. (…) For the same configuration of located object and reference object, when the reference object is labelled a plate, *on* is judged appropriate, but when the reference object is labelled a dish, *in* becomes appropriate. So how the same object is conceptualised can influence the types of computations that are performed on the visual scene."

It is assumed here that spatial terms do not involve a very high number of associated functional features, defined by the requirement that they hold across discourse contexts and are therefore part of the semantics of the terms. In contrast, there may be an infinite number of possible object functions and features, interacting with possible applications of spatial terms. While the object functions themselves may not be generalisable, the *interaction* with applicability may well be. Several proposals to this point have already been put forward in the literature. Since the amount of research in this field is vast (see Coventry & Garrod 2004b for an overview), I will restrict myself to the most relevant findings.

As a starting point I first consider some intriguing proposals put forward by Herskovits (1986). Consider the following famous example:

(81) The cat is under the table.

Here, the cat is clearly not under the lowest parts of the table – i.e., under the legs – but under the **functionally salient** part of the table, namely, the table top. Herskovits (1986:73) refers to this phenomenon as a **pragmatic principle of salience**:

> "One can use a noun which basically denotes a whole object to refer to the region occupied by a part of it that is typically salient."

This is not a random principle, but speakers must share some kind of experience in which the relevant part is salient. Such salience can, for instance, consist in the fact that only a part of the object is visible, as in *a rabbit under the bush.* This principle is well suited to account for findings such as the "toothbrush experiment" reported, for instance, in Carlson et al. (2003). Here, subjects were asked to place a tube of toothpaste above a toothbrush. Not surprisingly, many participants chose the functional part of the toothbrush rather than its center of mass for the placement. Thus, Herskovits' principle of salience does not only apply to object reference

(referring to the *table* instead of the *table top*) but also to interpretations of spatial relationships expressed linguistically. Spatial relationships are interpreted to pertain to the most salient or functional parts of the involved objects, rather than to their overall shape. The Attentional Vector Sum Model presented in Carlson et al. (2003) captures this idea on the basis of the empirical finding that speakers' **attention** is focused on the most salient object parts, i.e., those that people typically interact with.

A corollary of the pragmatic principle of salience is the following (Herskovits 1986:74):

> "The geometric description applicable may be the base of the object (i.e., its area of contact with the ground plane)."

This principle explains cases in which, of a three-dimensional scene, only two dimensions are relevant for reference, as in *the block in the circle.* This observation corresponds to the empirical finding that references to a *square* can be resolved unproblematically (though with increased reaction times) in situations involving only a cube (Duwe et al. 2002). However, according to Herskovits *two* dimensions cannot be ignored, using only one dimension, as would be the case in *the block in the line.*

In a similar vein, Pribbenow (1992:143) points out that certain regions may be prioritised for spatial descriptions. For example, the following two utterances both refer to the inside of a room, but still evoke different spatial representations:

(82) der Teppich im Zimmer
[the carpet in the room]

(83) das Gemälde im Zimmer
[the painting in the room]

In (82), the carpet is typically associated with the floor (thus confined to a two-dimensional representation of the room, rendering the vertical dimension irrelevant), while in (83), the picture is likely to hang on one of the walls. Thus, typical object placements evoke associations with different aspects of the spatial situation, which can of course be cancelled by providing a different kind of context (such as repositing a newly delivered carpet or painting in a room prior to arranging them at their future location). This may also mean that speakers associate specific **distances** with objects in a typical situation, as in:

(84) John was sitting in front of the TV.

Here, John is probably not sitting particularly close to the TV set but rather at a distance suitable for *watching* TV. Thus, the **region of interaction** (section 2.7) may be influenced by typical associations. Furthermore, the position of the referent may deviate completely from the geometric definition of a spatial term, if the objects involved evoke a purely functional interpretation, as in Klein's (1991:100) example:

(85) Unter dem Wams trug er ein rotes Leibchen.
[Under the jerkin he wore a red bodice.]

where the functional aspect of COVERING or OCCLUSION seems to be decisive.

Apart from the salience principles, Herskovits proposes that speakers also adhere to relevance principles (pp. 76ff.) which are closely connected to the functions of the objects involved. For instance, the examples

(86) There is some milk in the bowl.

(87) There is some dust on the bowl.

are distinguished not by the spatial relationship between the entities involved, as the choice of prepositions (*in* versus *on*) might suggest, but by the fact that in (86) the bowl's function as a container is important, while in (87), it is the bowl's surface that is relevant. Thus, in situations involving objects' functions a description is chosen that supports the association of

that function. This is accomplished straightforwardly by applying a term which contains a relevant functional aspect in its semantics.

With respect to empirical research in functional geometry, there is a large number of publications dealing, for example, with issues of containment in the application of *in*; see Coventry & Garrod (2004b) for an informative classification of findings. One main result of this research direction is that CONTAINMENT (or ENCLOSURE as used in Coventry & Garrod 2004a) as well as SUPPORT both reflect the effects of a "dynamic-kinematic extra-geometric routine of location control" (Coventry & Garrod 2004a:89) that can be – and has been – examined in detail by manipulating features and orientations of the objects involved in a spatial relation situation. The basic idea is that *in* and *on* are used together with containers and supporting surfaces, whose essential purpose is to constrain the location of other objects. The applicability of the prepositions then depends on the degree to which the objects involved can fulfill this function of location control. Therefore, the functional geometric relation is a matter of degree and can (in many situations) be transitive. It is specified in Coventry & Garrod (2004a) in terms of the degrees of connectivity expressed in the Region Connection Calculus (e.g., Randell et al. 1992).

With dimensional terms, results are less abundant; in fact, the only dimension that seems to have been researched in any detail is the vertical. Coventry & Garrod (2004a) propose that dynamic-kinematic routines similar to those observed for the application of *in* and *on* influence the applicability of the vertical prepositions: the effects of gravitation determine whether the relationship between referent and relatum is one of potential influence. These effects interact with the functionality of objects, such as the position of a coin slot on the back of a piggy bank (Carlson-Radvansky et al. 1999), which influences the application of *above.* Furthermore, experiments by Carlson-Radvansky and Irwin (1993) show that the applicability of *above* with an intrinsic interpretation (which is usually less acceptable than a gravitation-based interpretation, cf. section 2.2.1.5) increases if the salience of the relationship between referent and relatum is enhanced, e.g., by adding another reference object, by using only intrinsic descriptions, and by shortening the distance between the two objects. However, several results indicate that *above* is not as much influenced by functional factors as is *over*. For instance, Coventry et al. (2001) found differences in applicability between the different vertical expressions *over / under* vs. *above / below* when used in situations involving an object typically associated with a protective function. *Over / under* were more suitable to support this associated function than the other prepositions. Similarly, Coventry & Mather (2002) present the results of several experiments in which the applicability region of *over* was highly influenced by the given context (for example, a plane dropping a bomb on a city), while *above* was not affected in this way. This is in accord with the proposal of Tyler and Evans (2003) who suggest that the semantics of *over* and *under* contains a functional element of potential reach, in contrast to *above / below* which are not associated such a functional element (see p 12 above). Thus, the objects' functionality influences the area of applicability only when terms sensitive to this functionality are used, which seems to be a matter of degree.

Across axes, as described in some detail in section 2.2.1.4 above, Carlson-Radvansky & Radvansky (1996) as well as Taylor et al. (2000) found that intrinsic reference frames are preferred in situations involving functional features of (familiar) objects, while relative reference frames are more likely in non-functional situations. This observation is in accord with the findings on (functionally defined) proximity summarized in section 2.7 above, but it also extends them by suggesting that relative reference frames may be more flexible in this regard. However, as Taylor et al. (2001) show, this finding does not extend to unfamiliar objects.

That the functional relationship in interaction situations may be decisive is obvious considering a thought experiment: Imagine a situation where a person is seated at some

distance from a television set. The degree to which *in front of* is applicable will vary depending on two factors, namely, the distance between person and TV, and the functionality of the TV set. If the set is known to be out of order or not in operation, applicability should decrease with increasing distance faster than if the set is currently in use. Moreover, if the person is assumed to be trying to repair the apparatus, *in front of* will probably only be applicable in a fairly limited area close to the set. Coventry & Garrod (2004a:116f.) report some evidence from a similar scenario with respect to the application of the distance-related term *near*.

In front of may represent a special case with regard to application in functional situations, because the frontal axis is privileged in some respects (section 2.1). For example, interaction between two human beings usually occurs in a "canonical encounter" situation (Clark 1973) where the two persons are oriented towards each other. Likewise, interaction between an object and a person is most likely when facing the object, which itself is oriented with its functionally defined front side towards a person. Accordingly, Tyler & Evans (2003) propose a functional aspect of ACCESSIBILITY for *in front of*. Additionally, *in front of* is the only relation that allows for the usage of intrinsic reference frames induced by the features of the referent rather than the relatum (cf. example (15) above). Therefore, it would be interesting to investigate more details, comparing the axes, in the style of Carlson-Radvansky & Radvansky's (1996) experiment where differences between axes are unfortunately not spelled out (see section 2.2.1.4 above).

From this kind of evidence regarding topological terms, dimensional expressions, as well as some further findings on distance-related terms and *between*, Coventry & Garrod (2004a) develop a functional-geometric framework that offers an account of how these spatial terms are interpreted in natural discourse by way of establishing **situation models** (p 146):

> "[V]arious constraints work together to determine the appropriate meaning of a spatial expression in context. (…) [S]ituation models offer the vehicles by which these multiple constraints come together. Meaning in the functional geometric framework reflects the result of all these constraints coming together in the situation model, which supports the most informative relation between a reference object and a located object. By informative, we mean that the model should support the strongest inferences that could be drawn about the scene."

Such inferences can include situation-specific assumptions about why the speaker chooses one spatial term in favour of another, as in example (88) (cf. Coventry & Garrod 2004a:145) where the spatial relationship may be correctly described by *near*, but *at* would be a more usual description if the man is indeed playing. The authors suggest that (88) can be used in ironic contexts, implying that the man is not a good pianist.

(88) The man is near the piano.

Thus, the choice of a spatial term reflects underlying assumptions about the relationship between referent and relatum, not only with respect to the geometric relation, but also with respect to additional factors that may be context-related, due to dynamic-kinematic aspects, and object-function related. On encountering an assignment of a spatial expression to a relation, the hearer can make suitable inferences according to the discourse and situational context.

There is some evidence that another term is even more sensitive to functional aspects, namely, the primarily temporal term *before* which can also be used in spatial situations, although it is not used frequently in everyday discourse. *Before* can be used in a locational sense primarily if a functional aspect of potential interaction is associated. McIntyre (2001) contrasts the following examples, which – interestingly – are very similar to example (88) above presented by Coventry & Garrod (2004a:145) with respect to the difference between *near* versus *at*:

(89) Gertrude was sitting before the piano.

(90) Everyone in the room was fixated on the television. Fran and Stan were on the sofa, Gordon was sitting near the fire and Gertrude was sitting {in front of/*before} the piano, her back resting on the keyboard.

According to McIntyre, (89) entails that Gertrude is looking at the piano, playing it, or otherwise interacting with it. If *before* is replaced by *in front of*, this kind of inference is rendered optional. However, a context in which interaction is excluded renders the application of *before* unacceptable, while *in front of* (which according to the findings reported above is also associated a functional component) is still fine. Thus, McIntyre (2001) claims that *before* in its spatial sense *requires* a functional interpretation, while *in front of* does not:

> "The interactional entailment does not come from the context in which *before* appears; rather the preposition insists that it be embedded in a context where interaction is possible."

This line of argument, while intuitively appealing, still awaits empirical foundation. Research on the functional components of prepositions as well as the contribution of knowledge about objects has largely focused on the more typical spatial expressions, usually those used for the vertical axis. But note that McIntyre's hypotheses are well compatible with the approach proposed by Coventry & Garrod (2004a).

3 Formalisations of dimensional relations

There are various attempts at formalising various portions of the above insights in terms of axiomatic and computational models, natural language processing systems, and so on. Discourse-related factors influencing application, as described in section 5 below, are usually not captured in these models (but see Porzel et al. 2002, cf. section 5.2.1, and Gorniak & Roy 2004, cf. section 0). In his general review of a broad range of models for spatial expressions, Mukerjee (1998) notices a shift in focus over time across approaches: While earlier models tried to capture the semantics of spatial expressions in terms of sharp boundaries and clearly defined – mutually exclusive – areas, more recent approaches acknowledge the "scruffy", contextually dependent and graded applicability of linguistic expressions.

Well-received approaches to formalizing the various topological relationships between regions are the already mentioned RCC approach (Randell et al. 1992) and the similar 9-intersection model proposed by Egenhofer & Herring (1991), which formalises binary topological spatial relations between areas, lines, and points. In Shariff et al. (1998), this latter model is refined to capture the semantics of natural-language spatial terms based on their geometry. Their empirical analysis of the applicability of a broad range of English natural language expressions reveals that topological aspects influence spatial-relation terms more than metric aspects do, that natural-language terms can often be applied in flexible ways, and that multiple terms may be used interchangeably to describe a spatial configuration. The analyzed terms in this work are mostly complex constructions, which can be applied, for example, to describe the spatial relationship between a region (such as a park) and a line (such as a street), e.g., *ends just outside* or *runs along boundary.*

In the area of dimensional terms, Eschenbach (2004) presents axioms (which are partly represented in informal terms in section 1) that combine geometric and functional aspects in the lexical entries. The functional aspects captured by Eschenbach concern the spatial relation (cf. section 1.3), not the functionality of the objects involved as discussed in section 2.8.

Hernández (1994) and Clementini et al. (1997) establish a unified framework for orientation and distance relations, based on the notion of **qualitative representations** as opposed to quantitative (metric) approaches:

> "Qualitative representations use discrete quantity spaces, where a particular distinction is introduced only if it is relevant to the context being modeled."

Rather than formalizing spatial templates as described in section 2.3, they propose employing various levels of granularity. With dimensional terms, this may mean that the space

surrounding the relatum is divided into as many discrete sections as needed in a specific situation. These sections are labeled by terms like *left* or *left front* and can be analysed with respect to their neighborhood relations. Rooted in the area of artificial intelligence rather than (psycho)linguistics, this formalisation does not attempt to model the application of natural language in discourse. Rather, models such as this are suitable for non-linguistic cognitively motivated computational approaches to spatial reasoning. Thus, it takes up ideas presented (from a more psycholinguistic viewpoint) in Retz-Schmidt (1988), and it accounts for a variety of reference systems. This reflects the fact (already mentioned above) that spatial reference systems are not solely confined to language; rather, language expresses underlying concepts of reference systems in specific ways. It is with respect to such underlying concepts, not with respect to application of linguistic terms, that spatial reasoning in AI is cognitively motivated.

Starting from Hernández' approach (which focuses on orientation), Moratz & Fischer (2000) present a model for generating and accepting spatial dimensional expressions in a robotic system. For interpretation, they assume overlapping half-planes as reference areas for spatial expressions. In generation, however, a non-overlapping sector model is used that employs disjoint quadrants (Figure 9). This differentiation guarantees that acceptance is more tolerant than generation, and that only those terms are generated that are in the focal area of the acceptance region. Objects are treated as point-like regardless of shape.

The main focus of this model is on the variability of reference systems and the identification of the intended object using any of the available systems in natural discourse. Three kinds of reference systems are initially represented in the model: intrinsic reference systems using the robot as origin and relatum, relative reference systems using a different object as relatum and those using the group of objects as relatum. In the latter case, the group centroid is computed as the abstract location of the relatum, which then serves as the basis for the computation of spatial relations. This model has been put to use in several human-robot experiments employing users unfamiliar with the system, and is being developed further, using more specified acceptance areas (Figure 10) and priority weighting (Tenbrink & Moratz 2003, Moratz & Tenbrink, subm.). In some cases, several different reference systems may simultaneously serve as basis for the spatial description that needs to be interpreted. In that case, each object present is assigned a graded truth value in each possible reference system according to the computational model. The result yields the most likely candidate for reference. In cases of several likely candidates, clarification questions need to be asked.

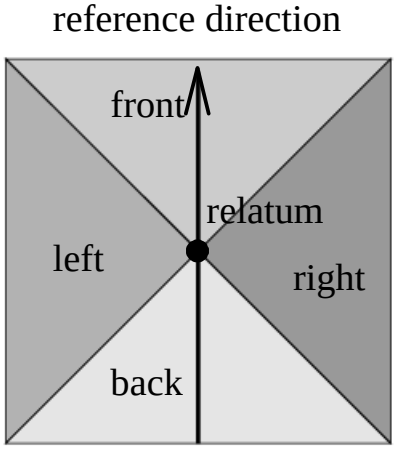


Figure 9. Non-overlapping sector model used in Moratz & Fischer (2000)

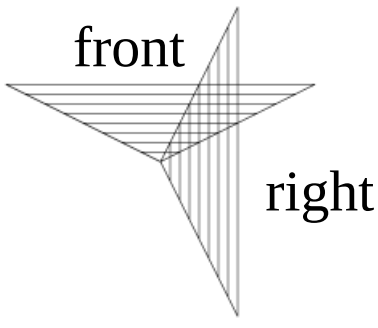


Figure 10. Overlapping acceptance areas (see Tenbrink & Moratz 2003)

The gradedness of likelihood of intended referents as modeled by Moratz represents one way of formalizing spatial templates (cf. section 2.3). These have also been subject to more direct,

detailed attempts at formalisation, for example, in Gapp (1994 a,b) as part of the DFG funded project VITRA (Visual Translator) (see also Retz-Schmidt 1988, Herzog 1995), Logan & Sadler (1996), and Vorwerg et al. (1997). In these approaches the aim is usually to model a typical spatial description of a referent with respect to a relatum, not (as in Moratz' approach) to identify an intended referent on the basis of a natural language description. The fundamental difference between these two kinds of discourse task is addressed in section 5.1 below.

Some approaches capture spatial relationships by way of vector models. For instance, Zwarts (1995) proposes a vector-based semantics for what he calls "relative positions" as expressed by a range of spatial terms, including dimensional and topological ones. Regions are analysed as sets of vectors. Thus, *above the house* is described as "the set of vectors that point from the house in an upward direction". Additional information about distance (such as *two meters behind the house*) can easily be integrated in this approach by restricting the length of the vectors. Zwarts explicitly abstracts away from different reference systems as well as object shapes, assuming that the relatum is idealized as a point.

In contrast, Regier and Carlson (2001) propose the "Attentional Vector Sum model" (AVS) that represents the influence of attention on specific parts of extended objects. The model starts by capturing geometric aspects such as proximity (cf. section 2.7), center of mass, and "grazing lines" (e.g., a horizontal line running through the topmost point of the relatum). Thus, they show how the shapes of extended objects influence the applicability of dimensional terms. Furthermore, the AVS model captures the fact that functional features of objects influence the regions of applicability, as described in section 2.8 above, by weighting the vectors by the amount of attention focusing on the relevant features of the relatum.

Another recently developed model capturing findings on functional aspects is presented in Coventry et al. (2005). The computational model implements the functional geometric framework presented in Coventry & Garrod (2004a) (cf. section 2.8 above) and is capable of assessing the appropriateness of dimensional terms for the vertical axis in a range of scenarios in natural discourse.

With respect to object shapes, Gapp (1994a:1394) proposes:

> "In most cases, it is sufficient to approximate the object to be localized with its center of gravity, since only position is required for the applicability of the spatial relation. In our system the following idealizations are used: (1) *Center of gravity*. (2) *Bounding rectangle (BR)* (…). (3) *2D representation*: The base of each object (necessary when perceiving objects from a bird's eye view, e.g., maps). (4) *Bounding right parallelepiped (BRP).* (5) *3D representation*: The complete geometrical description of an object."

Apart from the dimensional relations expressed by *in front of, behind, right, above,* and *below* Gapp's model also includes topological relations, *between*, and *beside*, which is interpreted as a junction of *left* and *right.* The model captures intrinsic and relative usages of the dimensional terms. In Gapp (1994b), an extension of the model is presented which also accounts for compositional relations such as *left and in front of* or a dimensional plus a topological relation such as *left and near.*

Gapp (1994a) assumes proximity as a requirement for applicability of all spatial expressions used; therefore, "if the distance from the LO to the REFO increases, then the degree of applicability decreases" (p 1396). This finding does not, however, correspond to the psycholinguistic results presented in section 2.3, which show that, within a visible area, the only effect of increased distance between the related objects is that greater distances from the focal axis become acceptable. The fact that the related objects also need to be proximal seems to be related to functional aspects rather than increase in distance as such (cf. section 2.7). In Gapp (1995) an empirical validation is presented where the implementation of proximity is dropped, reasoning that

> "the task of object localization, which demands the located object to be proximal to the reference object independent of the spatial relation, was separated from the task of apprehending a spatial relationship. The proposed model was therefore modified in that way that the distance which was multiplicatively combined with the angular deviation in Gapp (1994a) was dropped when computing the applicability of a projective relation."

Following this modification, the computational model could be successfully validated. The greatest deviations between the model's predictions and speakers' assessments appeared with the composite expressions, which is no surprise given the greater complexity and the availability of far less empirical evidence with respect to usage.

Within the DFG funded project SFB 360 "Situated Artificial Communicators", Vorwerg et al. (1997) present a computational model that places specific emphasis on the influence of object's shapes in 3D space. This is motivated by the project's aim of integrating visual, linguistic, sensory-motoric, and cognitive aspects in a system that can communicate with human users. In their approach, "surrounding boxes that are collinear to the objects' inertia axes as abstractions" (p. 159f.) are used. The approach captures diverging degrees of applicability, as obtained by recognition procedures and by psycholinguistic experiments. It integrates not only projective relations but also object type, colour, size, and shape. Information about spatial relations significantly improves identification of the intended object. Empirical evidence supports the validity of the implementation. See Vorwerg (2001:142f.) for a brief comparison of the model presented in Vorwerg et al. (1997) with other models, some of which have also been described in the present section.

4 Directionals

As already described in section 1.2, directional expressions are often viewed as similar (and secondary) to the corresponding locative terms. In the present work, the main focus is on locatives, since these are mostly used to spatially relate two or more objects to each other on a certain dimension. However, a few points can be made with respect to the interpretation of directional terms. Eschenbach (2004) proposes the following description:

> "The directional use of a preposition refers to a path that leads into a region as characterized by the locative use of the same preposition. Combinations of *nach* ('to') or *von* ('from') with one of the locative adverbs form directional adverbial expressions. (…) [T]he spatial condition expressed by the adverb (e.g., *oben*) specifies the goal region (*nach oben*) or the origin (*von oben*) of the path the composite expression refers to."

Thus, goal (or source) regions are defined in a similar way as regions in static situations. For instance, it is possible to define a goal (or source) region on the grounds of different reference systems, using an explicit relatum. In example (91) the origin of the underlying relative reference system (necessary because plates do not possess intrinsic sides) still needs to be identified, while in (92) the reference system is intrinsic and the relatum (the addressee) is identical to the origin.

(91) Put the cup behind the plate.

(92) Move the chair behind yourself.

But often, the relatum is not mentioned explicitly, as when an entity is moving in a direction specified by a directional:

(93) Now turn left.

Usually, in such utterances involving movement but no explicit relatum, the interpretation of the spatial expression depends on the intrinsic parts of the moving entity, similar to an intrinsic reference system in static situations. Thus, an utterance like (94) seems unlikely because it uses the observer's position rather than the moving entity's internal orientation for the assignment of direction.

(94) ? Move to the right from my point of view!

In contrast, (95) is unambiguously interpreted using the external regions as defined by the addressee's internal sides. The movement *to the right* is then a movement into the goal region on the right hand side of the addressee, as described by Eschenbach.

(95) Move to the right!

The situation becomes more complex if the moving object itself does not possess intrinsic orientations, as in a situation where an object is incapable of moving by itself but needs to be moved by external force. In that case, example (94) can be altered to a slightly more acceptable form:

(96) Move the ball to the right from my point of view!

However, this description is still unusual since speakers tend to use their interaction partner's perspective whenever action and movement is involved on the part of the listener (Herrmann & Grabowski 1994). Thus, a natural interpretation of (97) is that the listener's perspective is used since the object itself does not offer an intrinsic perspective.

(97) Move the ball to the right!

Here, the object is ascribed the same directions as the observer; then a motion to the right corresponds to the right half-axis of the observer.

Apart from the speaker's or interlocutor's perspective, it is also conceivable that a third entity defines the direction of movement of the object to be moved. This may be the case if this entity serves as background in the current situation, possibly because the moving object is located inside an object with intrinsic sides (e.g., in a train) or on its surface (e.g., on a desk possessing an intrinsic front because of features such as drawers). Thus, the following utterance can be interpreted unambiguously even though the interlocutors are not co-present with the object to be moved and can therefore not use their own orientation at the time of the utterance:

(98) When you get home, please move the lamp on the desk a bit more to the right.

Background objects may also influence the interpretation of movement directions with entities that themselves possess an intrinsic orientation, but only with expressions that can refer to the internal regions of background objects rather than (or as an alternative to) the direction of movement itself. This difference is exemplified by the difference in interpretation of expressions like *vorwärts / geradeaus* (forward / straight ahead) vs. *nach vorne* (to the front):

(99) Ich gehe nach vorne.
[I am going to the front.]

(100) Ich gehe vorwärts / geradeaus.
[I am going forward / straight ahead.]

If uttered on a train, example (99) would probably be interpreted to mean that the speaker intends to go towards the front section of the train, regardless of whether the speaker is currently oriented towards the train's front or happens to look towards a different direction. But example (100) can only mean a forward motion on the part of the speaker (defined by the speaker's orientation), which may or may not coincide with the forward direction of the train. Eschenbach (2004) notes:

> "The adverbs *vorwärts, rückwärts,* and *seitwärts* ('forward', 'backward', 'sideways') specify the alignment of a path relative to the intrinsic reference system of the figure. *Vorwärts* ('forward') expresses that the direction of motion is in accordance with the intrinsic orientation of the body. Thus, the reference system is bound to be intrinsic to the figure and cannot be specified differently by contextual influences. The geometric condition can be described as the alignment of the object order of the path with the intrinsic access order of the figure. The lexeme *rückwärts* is morphologically related to the noun *Rücken* (the body-part 'back') and *seitwärts* to the noun *Seite* ('side'). *Rückwärts* ('backward') expresses that the backside of the moving figure (…) is leading, i.e., precedes the center. Correspondingly, *seitwärts* ('sideways') can be used to say that a lateral side of the moving figure is leading."

The lateral axis does not offer such a distinction between only-intrinsic versus more flexible expressions in German, except for *seitwärts* which is unspecified for direction on the axis. In English, *leftward(s)* and *rightward(s)* seem to be available though used infrequently. The interpretation of (101) below, uttered on a train, would probably depend on the speaker's orientation as in (100), in spite of the fact that the surface form corresponds to that in (99). But this intuition may be due to the fact that the internal front and back regions of trains are much more prominent than their right and left sides. Thus, empirical investigations would be necessary to shed further light on these phenomena.

(101) Ich gehe nach rechts.
[I am going to the right.]

Furthermore, it is likely that the interpretation of *nach vorn* (to the front) is dependent on the availability and relevance of background entities with internal regions, such as the train in example (99). In other situations, a forward motion of the speaker may be more relevant, rendering the utterance synonymous to example (100). Furthermore, in cases where the interlocutors do not agree or are not sure about mutual agreement concerning the external reference system, the option of implicitly referring to an external region using *nach vorne* may not be available. In English, in contrast, it seems that a forward motion can only be expressed by *forward*, *straight (ahead)* and perhaps *ahead*, but not *to the front.*

Although the terms *forward / vorwärts* and *straight ahead / geradeaus* seem to be similar in some respects as shown in the discussion so far, there seems to be a difference with respect to applicability in *static* descriptions. Consider the following examples:

(102) Das Ziel ist geradeaus.
(103) The target is straight ahead.
(104) * Das Ziel ist vorwärts.
(105) *The target is forward.

Although it is possible in all cases to interpret the utterances with respect to a conceptualisation corresponding to "Move forward (in a straight line) in order to reach the goal", static descriptions seem only to be possible with *straight ahead / geradeaus.* This difference comes about by the fact that the latter two terms can linguistically relate to objects (i.e., nouns in a sentence), while *forward / vorwärts* always relate to a movement expressed in the verb:

(106) das Ziel geradeaus / *das Ziel vorwärts
(107) the target straight ahead / *the target forward
(108) Bewege dich geradeaus / vorwärts zum Ziel
(109) Move straight ahead / forward to the target
(110) * Bleibe geradeaus / vorwärts zum Ziel
(111) * Stay straight ahead / forward to the target

Schmidtke et al. (2003:167) distinguish projective prepositions and adverbs (i.e., those terms that are in the present work referred to as locative dimensional terms), directional adverbial phrases (e.g., *in das Haus* (into the house)*, nach rechts* (to the right)), rotational adverbs (e.g., *rechtsherum* (clockwise), and orientational adverbs (*vorwärts* (forward)*, rückwärts* (backward)), as well as some further adverbs that can be used to specify changes of orientation. Unfortunately, the terms *geradeaus / straight ahead* are not included in this categorisation. The authors point to systematic differences in the applicability together with certain types of verbs: for example, "orientational adverbs depend on the existence of a path of motion in the verb meaning" (p. 179). In contrast, they do not combine with verbs of orientation like *schauen* (look) which only point into a region defined by an external entity. Thus, while example (112) is unproblematic, (113) is not acceptable (Schmidtke et al. 2003:179).

(112) Paul schaut in das Haus. [Paul looks into the house.]

(113) * Paul schaut rückwärts. [Paul looks backwards.]

It seems, in contrast, that the directional terms *geradeaus / straight ahead* can be used with verbs of orientation (although this is not pointed out by the authors):

(114) Paul schaut geradeaus.

(115) Paul looks straight ahead.

Furthermore, considering the earlier discussion of the differences between *nach hinten* (to the back) and *rückwärts* (backward), (116) should be okay, since the region can be defined by Paul's intrinsic sides. *Rückwärts*, in contrast, can only be used together with motion.

(116) Paul schaut nach hinten.

[lit.: Paul looks to the back.]

These considerations suggest that there are systematic applicability conditions for directionals, which have – to my knowledge – not been investigated thoroughly using empirical evidence so far.

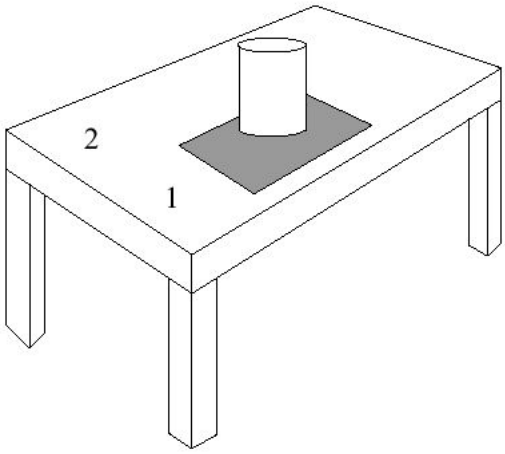


Figure 11. Setting in Jörding & Wachsmuth (2002)

One brief but interesting investigation was carried out by Jörding & Wachsmuth (2002) who asked 62 subjects to "move the coffee pot to the left" on a table that was oriented at a 45° angle with respect to their own position (see Figure 11). Since in this case the object to be moved (the coffee pot) does not possess unambiguous intrinsic orientations of its own, the directions must either be derived from the observer's position or from the background object (the table). It turned out that one-third of subjects interpreted the left direction with respect to the table (pos. 1) rather than with respect to their own actual orientation (pos. 2). The authors' interpretation of this result is that these subjects used the table's intrinsic sides for the definition of the movement direction. However, the reference system used cannot be viewed as genuinely intrinsic since the table (a simple one without drawers) does not possess intrinsic sides. Possibly, the desk pad on the table induces an intrinsic interpretation. Alternatively, the table's rather untypical orientation (i.e., its rotation of 45° relative to the observer) may have induced some of the subjects to imagine their own orientation *as if* they were standing directly in front of the table, choosing a typical position suitable for interaction close to their actual one. This would be a more natural position that accounts for the normal object function of the table, which is known to be relevant in spatial reference (cf. section 2.8). Then, the subjects could use their own intrinsic orientation to determine the left side of the table, which serves as background for the re-positioning of the coffee pot. This interpretation accords with the description of environment-induced reference systems in section 2.2.4 above.

Clearly, the interpretation of a directional in context may well depend on different pragmatic factors than the interpretation of a locational term. For instance, while it is clear that a

forward motion in a standard case describes a motion at a zero degree angle with respect to the moving entity's orientation, in a context containing a path (such as a street with curves) it may need to be interpreted to mean something like *follow the path in a more-or-less forward direction.* Thus, Gryl et al. (2002:29f.) interpret the instruction *go straight ahead* as follows[13]:

> "This verbal expression indicates that VP is on a medium-object and must go on moving on this medium-object (…) in VP's front direction. (…) The exact direction of the considered medium-object is not important. It may be curved (…)."

In a different context, it may mean *you will find what you are looking for within the frontal area.* And if somebody who is already in a forward motion is addressed by *now to the right,* depending on context this might involve a motion toward, say, a 45° angle rather than 90°, since the forward motion is merged with the rightward motion. In a route instruction context, again, *turn to the left* induces a search for a path on the left hand side of the moving entity (Gryl et al. 2002:31):

> "This verbal expression indicates that VP is on a medium-object W_1 and must go to the first intersection of the medium-object W_1 and another medium-object W_m situated on the left of VP."

Furthermore, movements into a newly specified region need to be differentiated from rotational movements, which may not always be obvious, depending on context; a brief utterance like *to the right* or *rechts* may be intended to mean either of both. How rotational descriptions should be interpreted is, for instance, addressed in Habel (1999). Here, further difficulties with respect to the angle of rotation arise, which is also a matter of contextual influence.

Pragmatic factors influencing the interpretation of directionals like those hinted at here will not be further pursued here, since the main focus is on locationals. However, in contrast to some suggestions to the contrary it is assumed that the findings on locationals are transferable to the domain of directionals only to a limited extent; interpreting directionals in context will surely involve further unexpected complications.

5 Discourse-related factors influencing application

This chapter will address relevant findings in the area of discourse analysis with respect to their influence on the application of spatial dimensional terms. The discourse factors influencing linguistic choices broadly fall into three categories: discourse task, spatial strategies, and aspects of the discourse situation such as who is being talked to, and which mode (spoken or written) is being used.

5.1 Discourse task

Much research on spatial language focusses on the ways in which human speakers describe the spatial relation between two entities in a given context, answering a question like "Where is the object?" (e.g., the MPIP research reported in Levinson 2003, as well as many others of the publications mentioned in the previous sections). Since the focus of scientific investigation is predominantly on this kind of discourse task, the impression could arise that this is a representative kind of discourse – or even the only kind – where spatial language is systematically employed. But psycholinguistic experiments such as those designed for highlighting the graded structure of dimensional terms (cf. section 2.3 above) usually do not test for choices speakers make in actual discourse, which naturally depend on the level of specificity needed in a specific discourse situation. If a vague characterisation of the spatial relationship is sufficient, applicability areas for simple expressions might be much larger than in a different context where a precise description is vital (as in Zimmer et al. 1998, where the

[13] "VP" is short for "Virtual Pedestrian"; a "medium-object" is an externally deliminated path, such as a street, a walkway, etc.

interlocutor needs to find a hidden element on a screen), or where the "best fitting" expression is to be identified (Vorwerg 2001). Thus, the motivation why a spatial relation should be described at all (i.e., the **quaestio**) is decisive; as described by von Stutterheim et al. (1993:102) it inevitably influences the choice of linguistic strategies:

> "A quaestio determines both the structure and content of the answer-text provided: It introduces the state of affairs (…), a particular perspective on the situation (e.g., should a situation be described or narrated) and thereby criteria for information selection and referential movement. It introduces a particular topic-focus distribution with respect to the information belonging to the different conceptual domains (…). These constraints on the construction of a complex information structure affect the choice of linguistic form in representing the information."

In the area of spatial language, many different kinds of quaestios occur in actual discourse, resulting in diverging usages of spatial expressions. In route descriptions, for example, typically a goal location is described via reference to streets and landmarks which can be easily identified in the real world (e.g., Tversky & Lee 1998). There, salience and dimensions of buildings play a role in the choice of landmarks; spatial relations are often sufficiently outlined via simple and vague expressions. Such a scenario differs fundamentally from psycholinguistic spatial localisation experiments where participants are asked to specify an entity's location relative to another. The area of route descriptions is fairly well researched in the literature and will not be addressed specifically in the present report, since it involves a number of issues that do not directly pertain to the applicability of dimensional terms (but see Tenbrink, in prep).

A different discourse task in which spatial reference is employed concerns the description of spatial relationships in a given scenario, for example, the structure of an apartment (Linde & Labov 1975), or the interior of a room (Ehrich 1985). The strategies identified for such tasks will be briefly addressed in section 5.2.1 below.

Another fundamentally different quaestio from that usually addressed is that one of several similar objects needs to be identified on the basis of spatial location, due to the absence of other cues such as distinguishing features or perceptual salience. The relevant question in such a task is "Which" rather than "Where". I will call this kind of quaestio **referential identification**, thus pointing to the fact that this area of spatial language interacts with other kinds of reference to entities in the real world, on the basis of other kinds of features. The broader area of object reference is itself well researched in the literature; it will be dealt with in some detail below. Note, however, that while reference as a concept is also used to denote the resolution of reference items within a discourse, i.e., linguistic expressions (such as anaphora) that point back (or forward) to other linguistic expressions, the present section is only concerned with linguistic expressions that point to objects in the real world (exophora).

In many scenarios, language is not the only medium through which referential identification can be achieved. Much research deals with the question of how gestures are combined with language in order to facilitate communication (e.g., Rieser 2004, Sowa & Wachsmuth 2002). Gestures accompanying language also seem to enhance speakers' abilities at formulating complex spatial relationships (e.g., Emmorey & Casey 2002). However, in the present work I focus on situations in which only language is available for reference, since I am concerned with the application of spatial dimensional terms. It should be noted, however, that the interpretation of such terms may be influenced by accompanying gestures (e.g., Landragin & Romary 2003). Another interesting finding has recently been reported by Kühnlein & Stegmann (2003), using a scenario allowing for both gesture and spoken language in which an object is to be identified out of a large amount of objects. They found that verbal utterances tend to be shorter and less informative in a proximal area where eye gaze and pointing gestures are viewed as sufficient, while with increasing distance from the speaker, utterances become longer and more complex. If the distance to the speaker is too great to allow for object identification via gesture, the utterance is usually formulated in a way that is complex

enough to be sufficient for singling out the reference object. These findings fit well with previous findings on multimodality reported in Oviatt (1999), where similar effects of other modes of communication on the complexity of linguistic utterances are described, showing that multimodal language differs in several respects from unimodal language. Thus, it should be expected that the effects described in the present work cannot be generalised straightforwardly to situations in which other modes are available for reference. As pointed out by van der Sluis and Krahmer (2000:8f.), the interplay of verbal and nonverbal (i.e., pointing) reference is influenced by the effort required to use either mode in each specific situation:

> "[A] balance should be found between on the one hand the speaker's effort to produce a description and on the other hand the effort necessary for interpretation of this description by the hearer. Hence we assume that the decision to use a pointing act for distinguishing an object is determined by two factors: the effort of pointing and the effort required for a full, linguistic description. We assume that the effort of pointing is determined by two factors: the distance to and the size of the target object. (…) [T]he effort required for a purely linguistic description (…) is proportional to the number of attributes and relations needed to generate a distinguishing description. When the complexity of the linguistic description is above a certain threshold, the linguistic description generated so far is discarded and a pointing act is generated instead."

This description is well suited to explain the findings reported, for example by Kühnlein & Stegmann (2003), since it explains why speakers do not in all situations produce a detailed linguistic description, or consistently use pointing instead.

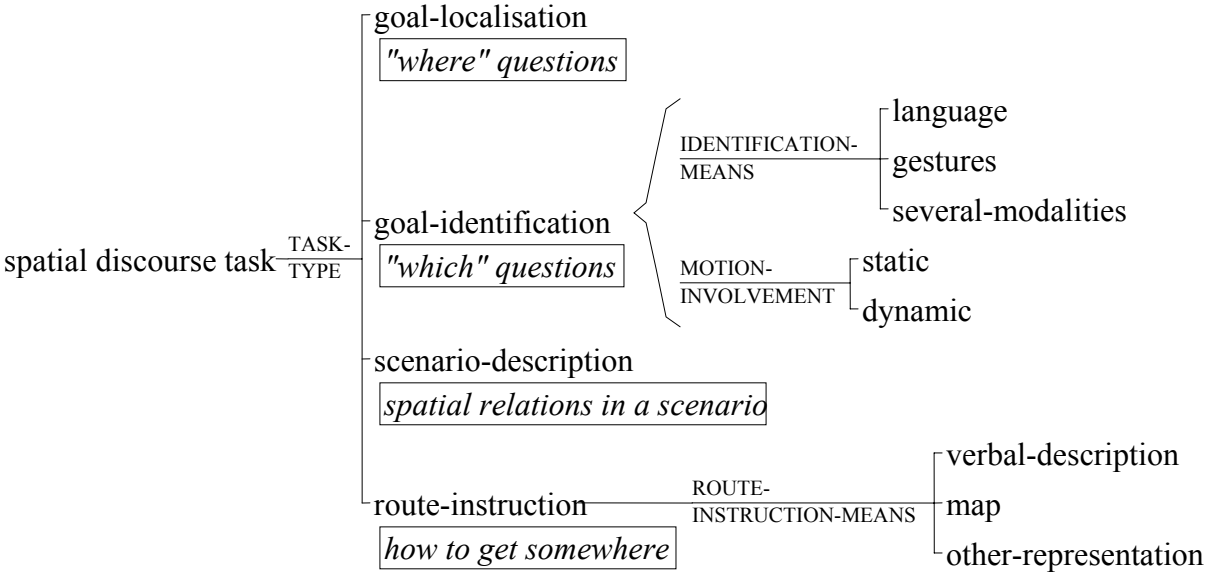


Figure 12. Task variability in spatial reference

Figure 12 summarises some ways in which spatial tasks can differ. Thus, the task may be to describe the position of a goal, i.e., an object is to be localised (as is the case when asking a "where" question). In a different case, a goal object may need to be identified out of a number of present objects (as is the case when asking a "which" question), or a scenario or a route may be described. Each discourse task may involve a number of suitable means for carrying it out. This variability is here spelled out only for route instruction and goal identification tasks. Not all means that are available in principle can sensibly be put to use in all kinds of task: it is simply not possible, for example, to describe the spatial relations of objects in a room using gestures, although gestures may accompany speech (as is the usual case in natural language).

Furthermore, goal identification tasks may be either static or dynamic: the task may be to simply determine its identity or to move to the identified object. In the dynamic case, instructions may become similar to route descriptions because speakers may opt to identify

the goal object by describing the route towards it. In both cases, instructions may be given either on-line or in advance, so that speakers' utterances may contain full descriptions or only parts of it. This kind of variability may either be the speaker's decision or it may be prescribed by the discourse situation. Such differences in speakers' strategies are a further issue for empirical work.

In the following section, I will specifically address issues pertaining to the task of goal identification using language only. This kind of discourse task is pursued further here because it has been largely neglected and is nevertheless crucial in the area of spatial reference.

5.1.1 Referential identification tasks

Identifying – or referring to – one of several objects in a spatial scenario differs in some basic respects from specifying a spatial relation. For instance, it should involve taking the overall configuration into account to a higher degree than in "Where" scenarios, in which other objects present can usually be ignored (or are simply absent in experimental settings). In contrast, "Which" questions presuppose (shared) knowledge of the situation to such a degree that the target object can be distinguished sufficiently from all other competing objects in the scenario, making the choice of spatial reference highly dependent on the number and arrangement of objects present.

Descriptions of spatial relations can also be used as a means of identifying – or referring to – a target object, such as when a "Where" question is posed in order to identify the correct target object out of several possibilities. Thus, describing a non-prototypical spatial relationship via a simple expression like *left* can be sufficient in a context where there are no competing objects in the left region of the relatum. An everyday example may illustrate this: "Where is the key?" can be answered sufficiently by a vague expression like "Left of the cup" no matter what the precise spatial relationship between the objects is, if there are no other keys closeby. Ultimately, the target object can be identified by establishing sufficient contrast to competing objects, similar to scenarios where "Which" questions are asked.

Linguistically, there are some systematic differences between utterances relating to the two kinds of discourse tasks. For example, the German dimensional adjectives (such as *das linke*) and the English dimensional superlatives (such as *leftmost*) can only be employed as an answer to "Which", not "Where" questions: There is no way in which "Where is the object?" can be answered by "The leftmost one"; even stating "It is the leftmost one" could at best be interpreted as a fairly indirect way of answering the question. Other syntactic options, in contrast, are available for both kinds of questions: "The object to the left of the barrel" (a typical answer to "Which object do you mean?") can easily be augmented to "The object is (located) to the left of the barrel" (a typical answer to "Where is the object?"). Note that German dimensional adjectives as well as English dimensional superlatives hardly occur in the literature on spatial language, mirroring the fact that discourse tasks of referential identification have been poorly researched so far.

In the broader area of object reference, a great number of issues and problems have been dealt with. Following Habel (1985:107) I will assume that reference concerns the pragmatic aspect of how (speaker's conceptions of) objects are referred to in real communicative situations, as opposed to the lexical semantics of the expressions that can be used for reference:

> "'Referenz' betrifft den pragmatischen Aspekt, d.h. die Beziehung, die ein Sprecher/Hörer zwischen sprachlichen Ausdrücken und der projizierten Welt herstellt. 'Designation' betrifft den semantischen Aspekt, d.h. die Beziehung zwischen sprachlichen Ausdrücken und projizierter Welt, die dem zugrundeliegenden Sprachsystem eigen ist."
>
> ['Reference' concerns the pragmatic aspect, i.e., the relationship that a speaker/hearer establishes between linguistic expressions and the projected world. 'Designation' concerns the semantic aspect, i.e., the relationship between linguistic expressions and projected world that belongs to the underlying linguistic system.]

Habel's notion of "projected world" accounts for the fact that speakers do not directly refer to objective reality but always to their conception of it; thus, language itself always has to do with speakers' projected views of reality, not with reality itself, whatever the differences may be. Evans (2004) elaborately advocates a similar, though perhaps more extreme view, pointing out (p 50) that

> "perceptual input has organisation imposed upon it which does not derive from the input itself; that is, the nature of conscious experience is structured by cognitive processes below the level of consciousness with the result that that of which we are directly aware, our conscious experience, does not necessarily equate with what may objectively be there. (…) I suggest that in so far as language 'refers', it refers to lexical concepts, the mental representations which may or may not reflect an objective world 'out there'. (…) Accordingly, language, and meaning, which it serves to express, must, on this view, be fundamentally conceptual in nature."

In the present work, the exact relation between speakers' concepts, perception, and objective reality is not specifically addressed. For present purposes, the most relevant insight is that language always represents different aspects of the world, reflecting the speakers' underlying conceptions. These may be influenced by subjective perceptions as well as underlying assumptions, and they may be altered throughout the course of interaction. Since these conceptions are crucial to the use of language (as observed by Habel, Evans, and many others), objective features of real world situations are less decisive than speakers' subjective views and representations of them. Thus, the "projected world" is assumed to be permanently present as an additional layer although this is not always spelled out explicitly.

Since I do not intend to cover the vast field of object reference exhaustively, I will focus on two aspects: First, I will briefly present Herrmann & Deutsch's (1976) account of psychological principles of object reference. Some of these, along with other assumptions, have been formalised for computational treatment in Dale & Reiter (1995), which is however not addressed in the present work. Second, I will try to identify issues that shed light on speakers' choices in which reference on the grounds of spatial position plays a major role. Such issues are explicitly ruled out in Herrmann & Deutsch's approach, though they do appear in subsequent work in this line of research (e.g., Pobel et al. 1988) as well as that following Dale & Reiter (e.g., Beun & Cremers 2001).

As is well known in the literature, in referring to objects speakers do not describe all features of the intended objects they can conceive of. Instead, they use linguistic representations that are specifically suitable in a discourse situation, for example, in order to distinguish the intended object from others present. In an account of how this is done, Herrmann & Deutsch (1976) formulated general principles of **greatest distance, dimension preference, redundant verbalisation,** and **partner-adapted verbalisation.** These principles capture that speakers, in choosing a reference strategy in an object identification task with several different objects, usually do at least the following:

First, they analyse the target object with respect to properties that can establish a (maximum) contrast to competing objects. Thus, if there are two black boxes of different size, the speaker chooses *size* for object reference. In case of several properties in which the objects differ, the speaker chooses the property where the distance to the competing object is most obvious. Thus, if there are two boxes, one of which is very small and dark blue, while the other is very big and black, *size* – rather than *colour* – will be chosen for reference (**greatest distance**). Individual preferences also play a role, especially if the distances are viewed as equal (**dimension preference**).

Second, the speaker encodes as many properties as needed for unambiguous object reference, but usually not more, being economic (cf. Grice's maxim of quantity, Grice 1975). But if the object reference task is complicated by the availability of multiple options, minimal differences in distance, and equal levels of dimension preference, speakers may encode more properties than needed (**redundant verbalisation**).

Third, speakers adapt to their interaction partner's view of the situation, taking into account cognitive and social distinctions, etc. (**partner-adapted verbalisation**).

Herrmann & Deutsch are exclusively concerned with object reference in non-spatial settings, designing their experiments purposively in a way that spatial reference is ruled out. In the present work, the main interest lies in situations where the opposite is the case, i.e., where target objects differ from competing objects only in spatial position, and additionally, objects of a different class may serve as possible landmarks (relata) for reference. This is a situation that has not often been directly addressed in the literature, since most research on spatial reference focusses on scenarios in which the position of an already agreed upon object is to be defined by relating it to another present object, usually of a different kind.

Previous experiments involving referential identification are reported, for example, in Schober (1993, 1998) and Mainwaring et al. (2003). In these studies, the main emphasis was on perspective choice; therefore, many of the aspects considered in the present work were not addressed specifically there. However, it should be noted that in Mainwaring et al. (2003) two further kinds of discourse tasks were examined: instead of achieving referential identification with an addressee, one task was to write a message to oneself, for later reference. Another task was to ask a question aimed at referential identification in a way that could be answered via a simple *yes* or *no* by the addressee. Results showed some differences with regard to perspective choice, with regard to choice of distance terms rather than dimensional expressions, and with regard to redundancy: offering an additional (alternative) description was more frequent when identifying the target object for an addressee or to the speaker's own future self, than in identification questions. Thus, the discourse task had an impact on speakers' linguistic choices in several respects.

As already described in section 2.2.1.4 above, Weiß et al. (1996) also present a study in which a dimensional term serves to identify a referent (in this case, a place rather than an object). Their results show that (at least German) speakers are influenced by situational factors such as formality and proximity, and that (at least English) speakers prefer non-dimensional expressions for referential identification when no oriented relatum is present, rendering intrinsic reference less accessible.

One recent study in which spatial referential identification plays a central role is presented in Gorniak & Roy (2004), who address speaker's strategies and range of variety in describing objects in a visual scene containing similar objects of two colours. Building on the results, they then present a visually-grounded language understanding model; thus, the main aim of their research is the implementation of results in a compositional parsing framework. Their implemented system is successful in understanding naturally produced object identification utterances. Their study was specifically designed to induce participants to produce spatial descriptions, because objects differed only in position (and colour as a non-discriminative feature). Participants were seated with their backs to each other, each looking at a computer screen displaying the same scene containing up to 30 objects, half of which were purple, the others green. One participant was asked to describe one object which the other was expected to identify and mark on their own screen, in which case the object disappeared. Thus, the number of objects in the scenario was gradually reduced. The collected descriptions therefore referred to a high number of different spatial scenarios with varying challenges and opportunities for referential identification. Since the listener was not allowed to speak, no dialogue developed, but the only feedback the speaker received was success or failure in identifying the intended object. Results show a number of insights gained in this intriguing kind of task. First, almost all utterances contain information about the non-discriminative feature colour. This way, the focus of attention is already canalised towards one half of the objects. Second, a very frequent strategy is "to refer to spatial extremes within groups of objects and to spatial regions in the scene" (p. 439). This is done, for example, by terms of

distance such as *closest*, or by dimensional terms such as *in the front* or *on the left side*. As indicated by the authors' interpretation of such terms as "spatial extremes", such expressions, though linguistically unmodified, refer to the object that is situated at the most extreme position as compared to other objects that may also be situated in the same spatial region, e.g., within the left half of the picture. Another possibility is to refer to the extremum using a dimensional superlative such as *leftmost.* Since the scenario used in the study is fairly complex, it is not surprising that such expressions are frequently used in combination with another description of a spatial region or extremum specification, as in *the frontmost left cone*, or together with other kinds of descriptions, such as grouping, as in *the rightmost green cone in the clump of four green cones to the right.* Grouping is generally used either for identifying a subpart of the scene in order to reduce complexity as in the previous example, or for using the group as a relatum, as in *the green cone at the back of the row of four.* Also single objects are sometimes (though infrequently, which is not surprising given the complexity of the scene) used as a relatum, as in *the green cone below the green cone.* Some utterances rely on the previous discourse in that they establish a spatial relationship to the previously removed object, as in *the green one right behind that one*. Apart from these main strategies, there is a high number of further, individual strategies (13% of the data) that were not covered by the previous classifications. This result shows the considerable degree of flexibility and creativity in speakers' spontaneous spatial descriptions in complex scenarios.

Since the objective of Gorniak & Roy's study was not predominantly linguistic but rather technologically motivated, the major focus of the analysis lies on the requirements for a computational system capable of parsing the kinds of utterances produced by the participants (see also Roy 2002 for an account of earlier work in this direction, based on machine learning). The degree of insight covered in this work is remarkable; it is a promising future goal to integrate this approach with related recent endeavours, such as those reported here and the empirical work presented in Tenbrink (2005).

One study in which the contrastive principles established by Herrmann & Deutsch were applied and investigated in a scenario allowing for both spatial and non-spatial object contrasts is presented by Pobel et al. (1988). They discuss results in several different directions. One particular result is that speakers adjust their naming responses to the listener's visual preconditions of object identification: in a sub-part of the experiments, the speaker was told that the listener (imagined to be sitting in an adjacent room) looked at a black and white screen. In these cases, colour was not mentioned by the speakers even though it was the major discriminative dimension. Usually, in the case of multiple verbal codability of the target object speakers preferred the visually most salient object feature for naming. Another interesting factor in Pobel et al.'s study is the fact that, in one part of the experiments, a listener was personally introduced to the speakers, while the other speakers were simply told that there was a listener. In both cases, however, there was no interaction possible, i.e., no feedback from the listener. The results show no difference between the two conditions.

Reference on the basis of *spatial* contrast was explicitly excluded in one part of the experiments, while the other participants were implicitly allowed to use it. The configurations consisted of 24 objects each, which were arranged in rows. Speakers allowed to use spatial reference employed it in 75.9% of cases, often together with other features. The authors formulate a caution with regard to this finding (Pobel et al. 1988:35):

> "Die matrixförmige Anordnung des Materials legt die Verwendung räumlicher Positionsangaben nahe, so dass aus der Auftretenshäufigkeit solcher Benennungen in dieser Untersuchung nicht unmittelbar auf die generelle Relevanz des Merkmals Raumposition für die Benennung von Objekten geschlossen werden kann."
>
> [The matrix-shaped configuration of the material suggests the employment of spatial position descriptions, so that the frequency of occurrence of such descriptions in this study does not

> directly allow for inferences about the general relevance of the feature spatial position for the reference to objects.]

Also the high number of competing objects, together with the clear spatial arrangement, poses a highly specific situation which is probably not generalisable to other situations. Thus, it is not clear whether speakers generally prefer to use spatial contrast in favour of other kinds of features that distinguish a number of competing objects.

Another caveat mentioned by the authors is that the speakers consistently developed and adhered to a range of different verbalisation strategies. The authors interpret this finding as an artifact of the experimental situation and suggest developing more natural scenarios in which such individual strategies are discouraged. However, this finding can also be interpreted in a more positive way, highlighting speakers' intuitive strategies in solving communicative tasks. Such strategies are especially interesting, for example, in human-robot interaction scenarios in which speakers often draw upon previous experience of success or failure.

Altogether, the results presented by Pobel et al. already suggest that the principles established by Herrmann & Deutsch can be applied also to spatial scenarios. However, due to the matrix-like arrangement of a high number of objects the only kind of spatial reference found in their study relied on rows and columns. This leaves us with the question what kinds of principles apply in the usage of dimensional terms. The main research question to be examined is:

What principles of object reference apply when only spatial reference is available?

Herskovits (1986) noted that, although the graded structure of dimensional terms applies in most communicative contexts, there are situations in which an expression like *to the right*, without modification, is capable of denoting the full right side (i.e., a half-plane) with respect to the relatum (p. 182):

> "[T]he loosest interpretation of the preposition is adequate, provided that obvious contrasts in the context allow the expression to fulfill its function of identifying the place of the located object."

This applies if there are no competing objects in the same spatial region. In case of the presence of further objects within the half-plane, unmodified dimensional terms can nevertheless be employed for contrastive reference. In that case, Herskovits' **shifting contrast near principle** applies (p. 81):

> "If two objects, A and B, are placed in a relation to a reference object in such a way that the ideal meaning of a preposition (…) is truer of A than of B, then one can use that preposition to discriminate A from B so that the locative phrase will be assumed true of A but not of B."

For instance, if *to the left* is truer (i.e., more valid) of A than of B, i.e., A is closer to the left reference axis than B is, A will be recognized as the target object even if *to the left* could also be applied to B. Here, gradedness effects again come into play, in that the object closest to the prototypical direction will be identified (Freksa 1999).

Another indication of what happens when several objects are located within one spatial region comes from the evidence reported above by Gorniak & Roy (2004), who interpret utterances like *on the left side* as referring to the object that is situated at a more extreme position to the left than any other objects in the scene. In this case, not the relation to the prototypical axis is relevant but the relative position on that axis, as compared to the competing objects.

Thus, in a discourse context involving referential identification (in contrast to spatial location) the immediate spatial setting supplies the information needed to interpret the utterances. Freksa (1981) elaborates on this point in a computational framework intended to deal with imprecise knowledge. According to his view, natural language descriptions provide ideas that (1981:2)

> "are not complete or crisp enough to enable the reader to reconstruct the described objects in such a way that they would closely resemble their models. However, they may be good enough to enable the reader to identify the objects in a given situation. The environment supplies

> information missing from the description to make correct identification of the target object possible."

Thus, the information represented linguistically may be incomplete and vague, and need to be supplemented by further information from the spatial situation, world knowledge, and inferences.

The process of identifying a reference object on the basis of a (contrastive) description is further complicated by different ways of interpreting linguistic expressions between the interlocutors. Freksa (1980:8) models the identification process as follows:

> "The context data base is searched first for objects whose features correspond to the highest possibility values associated with the description. The possibility distributions which are relevant here are the ones which the interpreter associates with the descriptors rather than the distributions of the describer. If no perfect match between description and object is found, objects with lower compatibility between the interpreter's meaning of the description and the actual features are searched for."

Joint reference can only be achieved successfully if the interlocutors' interpretative processes, which interact with their perceptual abilities, can be aligned sufficiently. This is true for all kinds of communication situations but may be specifically complicated in a situation where the interlocutor's perceptual as well as linguistic capabilities are unknown to the speaker.

It is a well-known general fact about language usage (e.g., van Deemter & Peters 1996, Carston 2002) that utterances can be underdetermined and implicit in a number of respects. In a spatial context, the immediate real-world situation is specifically decisive. In order to distinguish the inferred object from competing ones, it must first be decided which objects within an environment are relevant, i.e., whether they belong in the domain of reference or not, which may or may not be a matter of subjective interpretation in a discourse context. For instance, in a context of human-robot interaction the user may need to decide which objects will be recognized by the automatic system in order to serve as possible distractors or competing objects for reference.

Furthermore, not only relevance but also focus of attention is important (Grosz and Sidner 1986). There is a fair amount of research showing that object reference (in formulation as well as interpretation) is highly influenced by the focus of attention with regard to a subset of objects present (e.g, Beun & Cremers 2001, Regier et al. 2004). Such research often deals with spatial as well as other kinds of features which are available for reference. For example, results of empirical work in the project "Reference in Discourse" (part of SFB 360, Bielefeld) highlight regularities in the interpretation of descriptions that facilitate successful communication (e.g., Kessler et al. 1996). In ambiguous situations, reference is resolved with respect to a subset in focus. Note that this is a second way in which attention plays a role in spatial reference, in addition to the attention focus on functional features as modelled, for instance, in the Attentional Vector Sum model (Regier & Carlson 2001, cf. section 3). A third way is the establishment of focus and foregrounding on a purely linguistic level. All three have similar effects of influencing the interpretation of spatial descriptions. However, they may differ in importance in the presence of more than one factor, as van der Sluis and Krahmer (2000:11) point out:

> "We assume that linguistic context salience is primary, for instance, in the sense that an object *r* which was just described is more salient than an object which is in the current focus space (i.e., close to *r*) but has not itself been mentioned so far. In a similar vein, we take it that an object which is in focus is somewhat more salient than an object which is inherently salient but falls outside of the current focus space."

Furthermore, according to Kessler et al. some prominent features of other objects close by, such as colour, affect interpretation. Such generalisations have been implemented in a model of reference resolution called OINC (Object Identification in Natural Communicators), as described in Duwe et al. (2002).

Similarly, Brown-Schmidt & Tanenhaus (2003) show that referential domains constrain both the form of referring expressions, and their interpretation. According to their empirical work, factors affecting the mechanisms of reference interpretation in interactive conversation include explicit mention of referential domain, in which case there is no interference from potentially competing referents outside the domain, the salience of objects, presence of competing objects with similar phonetic sequence, and the development of interactive dialogue strategies between participants. The latter factor concerns the (by now) well-established finding (e.g., Clark & Wilkes-Gibbs 1986, Pickering & Garrod 2004) that reference is often achieved as an interactive process. For instance, Rieser (1996) shows that conversational participants often agree on situationally dependent representations, such as metonymies, in order to achieve smooth and effective communication (see section 5.2.1 below). Moreover, especially in cases of qualitative or vague spatial instructions, interactive negotiation supports the decision about how the instruction should be interpreted (Wachsmuth & Cao, 1995). Thus, spatial referential identification can be described as a speaker's decision as to how much, and which, information to convey on the basis of the external situation as well as the dialogue history and the interaction partner.

On the basis such well-founded findings in the literature, Beun & Cremers (2001) formulate a number of hypotheses with respect to object reference in a multimodal setting. Their starting point is the assumption that speakers use minimal effort in conversation (cf. the *principle of minimal cooperative effort* formulated by Clark and Wilkes-Gibbs 1986). The first hypothesis postulates a potential for ambiguity in speakers' utterances on the basis of salience:

> "1. If the target object is inherently salient within the domain of conversation, use reduced information."

This hypothesis, though following indirectly from the findings reported above, has not yet been very well empirically proven. At the present stage of research, it seems that there is at least some tendency towards reduced complexity within focus spaces (see e.g. van der Sluis and Krahmer 2000). The next hypothesis addresses contrastive reference:

> "2. If the target object is located in the current focus area, use only information that distinguishes the object from other objects in the focus area."

This hypothesis should be directly transferable to a setting in which only spatial reference is available for singling out the target object, as targeted in the present work. The following hypothesis, in contrast, concerns speakers' references to kinds of object features:

> "3. Use absolute features as much as possible and use relative features only if necessary."

This hypothesis is based on the observation that using absolute features is easier because no comparison to other objects is necessary, as is the case with relative features. Thus, it is assumed that absolute features may be mentioned even if they are not discriminative, but relative features are added if necessary for ruling out ambiguity. Notice that this assumption is already implemented in Dale and Reiter's (1995) algorithm (mentioned above), where it is also assumed that the object *type* (a most obvious object feature) should always be mentioned, irrespective of its discriminativity.

Finally, a different topic is addressed by the fourth hypothesis, namely, the choice of relatum:

> "4. If an explicit relatum is needed for referring to the target object, choose as relatum an object that is in the focus of attention."

This again points back to the importance of focus areas in complex scenarios, which are thus shown to be relevant in several different respects.

Starting from these observations we are now in a position to formulate more concrete research questions – and tentative answers – for the achievement of referential identification in a situation that only allows for *spatial* reference in order to avoid ambiguities.

How do speakers choose a reference system and point of view? If a partner is present it can be expected that participants will often choose their partner's perspective, especially if the

partner is expected to act (Herrmann & Grabowski 1994:123), and that speakers will adapt their utterances to their interlocutor's in various respects (Clark 1996; Pickering & Garrod 2004), for example with respect to the choice of reference systems (Watson et al. 2004). These findings further specify Herrmann and Deutsch's principle of partner-adapted verbalisation (see above). But the literature does not provide much evidence with respect to which reference systems are preferred if several options are available in an object identification scenario. This is so in part because such scenarios have not been in focus very often in spatial cognition research, and specifically, group-based reference has hardly been mentioned in the literature so far at all. Likewise, little can be said for the case of several options for perspective when there is no interaction partner. In the present work it is assumed that the identification task – requiring reference on the basis of spatial contrast – plays a role in the decision, since situations may arise in which one kind of reference system or perspective enables a clearer contrast than the other ones available. On the grounds of Herrmann & Deutsch's principles, this means that, *in addition to adapting to their interlocutor, speakers choose a reference system and perspective that is suitable* (just as a unique object property is) *for distinguishing the target object from competing ones within the referential domain*. If attention is focussed on specific subsets of objects in a scenario, a reference system and perspective may be chosen that only distinguishes objects within that subset.

How do speakers choose a reference axis? Within a reference system, the frontal (*front/back*), the lateral (*left/right*), and the vertical (*above/below*) axes are available for reference. With spatial localisations, the reference axis is chosen that the target object is closest to. But with referential identification, this may not yield unambiguous reference. It is hypothesized here that competing objects play a role in deciding about a reference axis: *A reference axis is chosen that is best suitable for distinguishing the target object from competing ones within the referential domain and/or the subset in focus, considering the principle of greatest distance*.

How explicit are speakers about underlying reference systems and origins? The situation may offer various reference systems yielding similar results (e.g., Carlson 1999): in some situations, *to the left* could equally well be used for an intrinsic reference system conflating origin and relatum, for a group-based reference system using the other objects as relatum, and for a landmark-based relative reference system. In such situations, *to the left* can simply be used without further specification since no conflicts arise. Similar observations apply with respect to the chosen point of view (origin); if several options are available yielding no difference people do not need to provide an expression like *from my point of view*. In intrinsic reference systems, the origin is often specified because it coincides with the relatum, as in *in front of me*. The interesting case is when different reference systems and points of view yield different results, so that *to the left* can be interpreted in different ways. Herrmann & Grabowski (1994:132) state that speakers are usually *not* explicit with regard to the perspective used. This may be because they tacitly assume that the interlocutor will understand the intended meaning even without the additional effort, since perspective needs to be expressed by additional linguistic material and since the partner's perspective is (in many kinds of situations) conventionally preferred. But with respect to reference systems, no such conventions are known so far, so that it is hypothesized here that *relata – but not necessarily origins – will be made explicit in case of potential conflict*.

Under what circumstances do speakers modify and combine dimensional terms? In spatial localisation tasks, speakers increasingly use modifiers and compounds as distance to the reference axis increases. According to Herskovits (1986) this is not expected in identification tasks (see above). Likewise, Herrmann & Deutsch's principles, as well as Clark & Wilkes-Gibb's principle of minimal cooperative effort, predict that speakers will not provide more information than needed, unless several options with equal properties compete. Thus, it is hypothesized that *speakers use a dimensional term without modification or combination with*

another dimensional term in case there are no competing objects for which the same description applies to the same degree. If the target object is placed where it could equally well, and equally unambiguously, be referred to by two terms, such as *to the left* and *in front*, both may be combined.

Thus, according to the findings presented in this section, it can be expected that pragmatic principles such as contrastivity can undermine prototypicality effects like those described in section 2.3 above. Crucially, using minimal effort for achieving contrastivity implies that no detailed spatial description will be given if this is not required by the scenario. This observation is a strong argument against including prototypicality structure in the semantic specification (cf. the discussion in section 1.3). Note, however, that the gradedness in applicability of dimensional terms is not completely ruled out in this case, since unmodified dimensional terms can only be used when there are no competing objects for which the same description applies to the same degree.

The validity of the above hypotheses needs to be addressed empirically, which is one of the aims of the project work in I1-[OntoSpace] of the SFB/TR8 on Spatial Cognition, which the present report is part of (see also Tenbrink 2005, Tenbrink in prep).

5.2 Strategies of describing spatial relations

In addition to – and in interplay with – the influence of the task given in a discourse, linguistic choices are also affected by another crucial factor, namely, by speakers' individual strategies. While a certain amount of idiosyncrasy must be expected, recent research has pointed to several systematic ways in which speakers can choose among available options for describing spatial relations. The most important of these concern, on the one hand, the level and framework of description, and on the other, the choice of Figure and Ground.

5.2.1 Level and framework of description

Several studies involving various kinds of tasks and settings have shed light on speakers' strategies in describing spatial relations. Very often, such strategies can be shown to be interrelated with the listeners' reactions, i.e., are an outcome of interactive collaboration (e.g., Clark 1996). Here, the focus is on the range of variety in spatial strategies.

Langacker (1986:25f.) points out that spatial situations can either be conceptualised as sequential, or they can be represented as a whole, i.e., atemporally. He describes the distinction as follows:

> "I propose that the distinction between a process and a complex atemporal relation involves the contrast between 'sequential' and 'summary scanning'. Sequential scanning is the mode of processing we employ when watching a motion picture or observing a ball as it flies through the air. The successive states of the conceived event are activated serially (…). On the other hand, summary scanning is what we imply in mentally reconstructing the trajectory a ball has followed".

This difference explains the semantic closeness between verbs expressing spatiotemporal relations, such as *cross*, and prepositions expressing the same spatial relation in an atemporal way, such as *across.*

When using language to describe complex spatial situations, sequentiality inevitably comes into play because of the simple fact that linguistic representations are themselves linear, while space is three-dimensional. Thus, spatial information needs to be linearised. With spoken language, the sequentiality is realised through time, and with written language, through space. While temporal information can easily be presented linguistically in the same order as the events occurred (though this is not a constraint), with space there does not seem to be any default order. But in fact, there are some conventions with regard to how spatial information is presented. For example, if entities are conceived of as being ordered on one of the spatial

axes (usually the frontal one, see section 2.1) the order of mention may be pre-structured. A different situation is one in which a route is to be described. Then, linguistic descriptions correspond to the order in which the spatial entities on the route will be encountered when the route is actually travelled. Similarly, in instructions of tasks involving the construction of spatial objects, object parts will be mentioned in the order in which they are needed.

Experiments presented by Buhl et al. (2000) show furthermore that the order of representation is influenced by the order in which the objects in a spatial situation have originally been encountered, i.e., by the order of knowledge acquisition. The authors call this effect the **according-to-experience-effect** (**Geneseeffekt**). Moreover, if knowledge has been acquired in two diverging ways, the first of those is more decisive than the second (**anchor effect**). This effect was shown to be stronger than another well-known psychological effect called the **recency effect**, according to which the most recent experience is the most decisive one.

If the situation (or its original presentation in the process of acquisition) does not itself provide an order on any kind of axis, it is still possible to impose an order. Thus, Herrmann & Grabowski (1994:114) describe the option of describing spatial surroundings by way of conceptualising a **generic wanderer:**

> "Sprecher mobilisieren ihr *Handlungswissen*, um in ihr Sprechen über Raumkonstellationen ein geordnetes Nacheinander zu bringen. Dies geschieht so, dass sie dem Partner sagen, was man sieht, wenn man auf bestimmtem Wege durch die im Bewusstsein des Partners aufzubauende oder im Bewusstsein des Partners bereits vorhandene Raumkonstellation *hindurchgeht*. (Oder man beschreibt, was man sieht, wenn man lediglich den *Blick* wandern lässt.)"
>
> [Speakers activate their behavioral knowledge in order to sort their talk about spatial configurations into an ordered succession. This happens by telling the partner what they see when *walking along* a specific path through a spatial constellation which either needs to be established or is already present in the partner's consciousness. (Or they describe what can be seen by merely letting the *eye* wander.)]

This option is often realised in German by employing the indefinite third person pronoun *man* which denotes no specific person but rather an abstract personality (equivalent, perhaps, to English *they*). It is however also possible to address the interaction partner directly, and represent the situation linguistically (by using a second person pronoun) as though he or she did the "wandering" through the spatial situation.

A ground-breaking early study in this area was conducted by Linde & Labov (1975). They had participants describe the structure of their own apartments, and identified several systematic regularities in the participants' strategies. For example, they follow a mental route, beginning with one room, and systematically describing further rooms according to an imagined tour through the apartment. However, if a branch during the tour leads to a dead end, speakers mentally "jump" back to a previously described place. Speakers indicate directions using terms like *left* and *right*, and they differentiate between main and minor rooms.

Ehrich (1985) reports a study in which the interior of a room should be described, which was done either by employing the generic wanderer strategy, or by focussing on the spatial relations between objects in functional arrangements (such as tables and chairs). She notes that the former strategy represents a connective, linear description, while the latter may be disconnective, involving spatial leaps and discontinuities. In disconnective descriptions, the usage of dimensional terms is much more frequent than in connective descriptions, where the direction-neutral expressions (*neben*) and temporal terms (*danach*, describing the experiences of the imagined wanderer) are more likely to occur.

If the aim is to instruct the interaction partner to reach a certain position in space, a common strategy is to start by leading the partner into the target region using large spatial units in the description, then becoming more and more specific until the target is reached (Herrmann & Grabowski 1994:152):

> "Solche Lokalisationssequenzen bilden eine Grundstruktur unseres Suchhandelns ab: Jedes Suchen (…) verläuft sozusagen vom Groben zum Feinen."
>
> [Such localisation sequences represent a general structure of our search behavior: Each search (…) proceeds, so to speak, from the coarse to the detailed.]

Another strategy mentioned by Herrmann & Grabowski (1994) is to mention the category or framework of what is to be described. For example, after mentioning a *village* it is possible to refer to typical parts of villages by relying on world knowledge that can usually be safely assumed to be shared by both interlocutors (cf. the notion of Common Ground established in Clark 1996). The more is known about such activated frameworks, i.e., the more specific the framework offered by the speaker, the less complex the task of describing spatial relations may become. Such spatial hierarchical frameworks are often described in terms of conceptual models, which serve as basis for examining how humans deal with wayfinding tasks (e.g., Timpf et al. 1992).

Since this is so, speakers often activate frameworks that are not readily pre-defined in a situation. For example, Garrod & Sanford (1988) examine what they call "discourse models" (or situation models) which are constructed for the specific purposes of the present discourse, similar to the notion reviewed by Zwaan & Radvansky (1998). The basic idea is that situation models are required even when there is a real world scene to talk about, i.e., the conceptualisation of the scene plays a role in referring to it. Thus, in the experiments described in Garrod & Anderson (1987), speakers developed consistent description schemes for reference to objects in a maze, a process which Garrod & Sanford (1988) claim to be dependent on underlying mental models of the maze configuration (1988:149f.):

> "Thus mental models of space have the effect of breaking down any scene into significant spatial entities – points, lines, regions or volumes of space – associated with the various objects in the scene, and then representing significant spatial relations between those entities".

The imposed relations between the entities in the maze game then correspond to other kinds of relations perceived between objects in the real world, e.g., the functional relationships described in section 2.8 above. The analysis shows that the underlying models can constrain or enhance referability and discriminability of entities in the maze situation.

Rieser (1996) shows that speakers metonymically refer to a *plane* in a task where a toy model of a plane is to be built (via spoken instructions) out of several ("Baufix") elements that partly do and partly do not represent specific modules of a real plane:

> "Bei der globalen Strategie benennt der Instruktor das ganze Objekt anfänglich als *Flugzeug, Propellerflugzeug* oder (fälschlicherweise) als *Helikopter*. Dadurch wird eine Art von *top-down*-Information gegeben und die Repräsentations-Metonymie kann, falls die Koordination der Agenten klappt, auf Aggregate 'heruntergedrückt' werden. So ergeben sich *Höhenleitwerk, Cockpit, Motor* oder *Propeller* quasi von selbst. (…) Repräsentations-Metonymien werden nicht ausgedehnt auf Objekte/Eigenschaften, die im 'Baufix'-Modell rein instrumentell oder arbiträr sind. Rein instrumentell sind Schrauben, Muttern oder Löcher. Sie halten das 'Baufix'-Aggregat zusammen."
>
> [In the global strategy, the instructor first refers to the whole object as *plane, propeller-driven plane*, or (erroneously) as *helicopter.* Thereby a kind of *top-down* information is provided, and the representational metonymy may, if the coordination of the agents is successful, be transferred to aggregates. Then, *elevators, cockpit, motor*, or *propeller* can easily be referred to. (…) Representational metonymies are not transferred to objects or features which in the 'Baufix' model are purely instrumental or arbitrary. Bolts, nuts, and holes are purely instrumental. They hold the 'Baufix' unit together.]

Rieser notes that this strategy turned out to be quite effective, clearly enhancing the referability of objects and pieces. For example, by employing the plane metonymy the speaker can already very early point out where the intrinsic regions (e.g., the *back* and *front* parts) will be. Then, in later instructions they can use these regions for reference.

This option is not available in a different kind of strategy also employed by some of Rieser's subjects, which he calls the **local** strategy. Here, the overall framework of a plane is not referred to, and speakers struggle with local descriptions of single parts:

> "Man kommt auch mit der lokalen Strategie durch. Sie ist allerdings sehr viel weniger effizient als die globale, da sie Probleme bei der Verwendung von Positions- und Richtungsangaben schafft. Der Koordinationsaufwand für die Agenten ist bei der lokalen Strategie weitaus höher als bei der globalen, d.h. sie müssen mehr Zeit auf Tests, 'Backtrackings' und die Sicherung des erzielten Ergebnisses verwenden."
>
> [Local strategies may also be successful. However, they are much less efficient than global strategies because they create problems in the descriptions of positions and directions. The agents' coordination effort is much higher with local strategies than with global ones, i.e., they need to invest much more time for tests, backtracking, and the securing of the obtained result.]

Local versus global strategies have also been identified by other authors in different kinds of tasks. For example, Carroll (1993) distinguishes three categories of describing the location of objects (1993:25), the third of which corresponds to Herrmann & Grabowski's (1994) generic wanderer:

> "(i) Global Frames of Reference. The frame of reference can be based on a single spatial structure which is applied so as to encompass the entity under description as a whole. (…) (ii) With a local, point by point system of organization, the spatial structures used are anchored at the individual parts which make up the entity under description. Parts and properties are described and located in terms of spatial structures anchored at this level and not on the basis of a global division into regions. Here parts of the entity are related successively to form a representation of the whole. This is termed an additive, point by point frame of reference.
>
> (iii) If the information to be expressed is organized in dynamic terms, the frame of reference can be based on what a person or fictive observer encounters on a tour of the entity or scene in question."

Carroll's empirical work showed that, in descriptions of an abstract L-shaped figure, many English speakers "ascribe a global shape which more or less matches intrinsic features of the entity (L-shaped, V-shaped, S-shaped, shape of a 4, etc.)" (p26), similar to the findings by Garrod & Sanford as well as Rieser. An implementation of Carroll's three-fold categorisation in natural language generation is presented in Porzel et al. (2002).

Figure 13 below sums up the different strategies for spatial reference in a schematic representation. It shows that global strategies may rely on two different kinds of frameworks: the interlocutors may rely on world knowledge in referring to object parts and directions based on an observed spatial structure (e.g., a village, cf. Herrmann & Grabowski 1994), or they may, as reported by Garrod & Sanford (1988), agree on a situation model for ease of reference within the discourse. Both options may also be intertwined within a discourse, as is true for all of the observed kinds of strategies.

The three different options of spatial descriptions are conspicuously reminiscent of Siegel & White's (1975) famous distinction between **landmark** vs. **route** vs. **survey** knowledge (assumed by the authors – but perhaps not in subsequent research – to be acquired by children in that order). Landmark knowledge can be viewed as similar to point-by-point or local descriptions, route knowledge mirrors the perspective of a generic wanderer, and survey knowledge may be linguistically expressed in terms of global frames of reference as described by Carroll.

The similarity of linguistic-pragmatic and psychological findings opens up a perspective on a vast research area in cognitive psychology, which can only be touched upon in the present work. A fair number of studies prove that knowledge about the position of single objects (or landmarks) differs fundamentally from dynamic knowledge about routes, and again, from static survey knowledge about configurations of objects. In large-scale environments, the latter represents knowledge similar to that represented in maps. But naturally, what is

represented linguistically does not necessarily correspond to the knowledge status of the speaker. Rather, different kinds of knowledge correspond to different kinds of linguistic strategy.

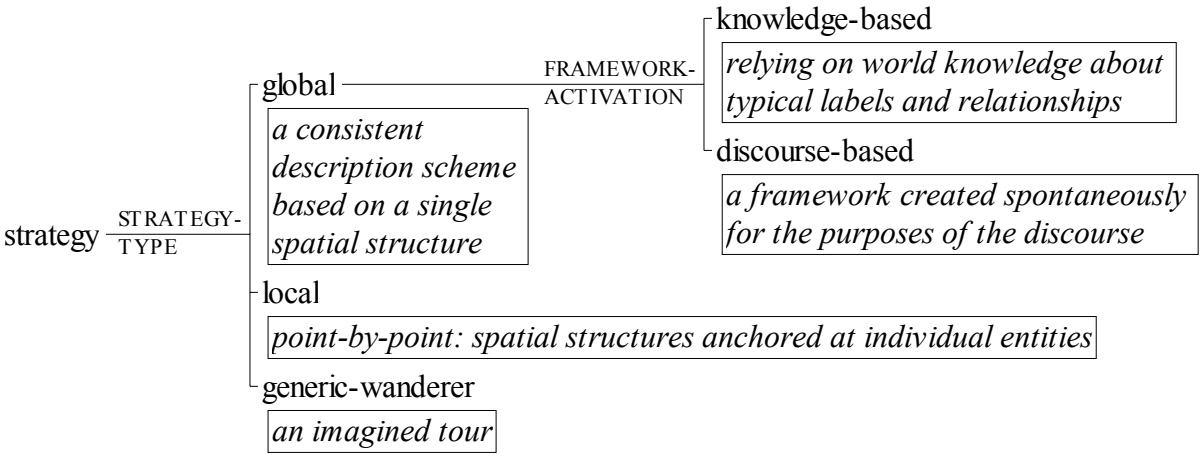


Figure 13. Strategies for spatial reference

There is not much literature on how the three kinds of knowledge interact with the three kinds of linguistic strategies. Taylor & Tversky (1992) show that mental models developed on the basis of text allow inferred spatial relations to be as accessible as explicitly mentioned ones. Taylor & Tversky (1996) investigate speakers' descriptions of sketch maps in terms of route versus survey perspectives, identifying different kinds of lexical choices associated with the different strategies. Herrmann & Grabowski (1994), who only distinguish between static versus dynamic localisation sequences, point out that not much is known about speakers' preferences in natural discourse (1994:144):

> "Es gibt schwache Anzeichen dafür, dass Sprecher eher dynamisch als statisch lokalisieren, wenn sie Kindern etwas erläutern wollen und wenn sie überhaupt andere Menschen zu instruieren beabsichtigen. Es sieht auch so aus, als ob eine bewegte Umwelt eher dazu einlädt, sie unter Verwendung von dynamischen Lokalisationssequenzen zu beschreiben. Nach unserem Eindruck sind diese potentiellen Effekte jedoch schwach."
>
> [There are some weak indications that speakers prefer localising dynamically instead of statically when they wish to explain something to children, or intend to instruct other people in general. It also seems that a dynamic environment generally invites description by using dynamic localisation sequences. According to our impressions such potential effects are however only weak.]

The effect of how the relevant spatial knowledge was acquired prior to the verbal description (the "Genese-Effekt") is much more important, according to findings reviewed by Herrmann & Grabowski. Furthermore, if the spatial knowledge is acquired in several different ways, the *earliest* experience turns out to be most relevant for the choice of strategies of linguistic description ("Anker-Effekt").

Similar effects are summarised in Taylor & Naylor (2002), who additionally show that not only the source of the spatial knowledge is important, but also the goal, i.e., why the knowledge was acquired (in this case, to learn the layout of a building versus to learn the fastest routes between locations).

One observation often pointed to is that speakers readily switch between different strategies within a discourse (e.g., Herrmann & Grabowski 1994, Carroll 1993). Tversky et al. (1999) address the question why this may be the case, in spite of the fact that this is cognitively costly. They find that adopting and maintaining a specific perspective (or strategy) also involves certain costs, which at times can be higher than switching to a different strategy that has become available (1999:404):

> "In some cases, the costs of switching perspective may not be greater than the cost of staying with the same perspective. Some spatial situations may be more readily described from one perspective rather than another. Indeed, some spatial arrays seem to encourage route perspectives, and others, survey perspectives. Having a single prominent pathway through an environment and landmarks at the same size scale, for example, encourages route descriptions".

Specifically dimensional terms are hard to compute, with the effect that "people avoid using them when other options are available" (ibid). This observation relates to findings reported in Mainwaring et al. (2003), who compare speakers' strategies of spatial reference in a referential identification task using a number of different configurations and scenarios. They find that speakers avoid using dimensional terms – specifically, if the lateral axis would be needed for reference – especially if the target object can be specified using distance terms. This is only the case, of course, if the competing object is not equidistant to the relatum (i.e., the entity to which the distance term refers, which can be the speaker, the addressee, or a landmark) as the target object. In that case, speakers prefer *near* to *far* even if this means that they specify the target object's relation relative to their own position rather than their addressee's, which is usually the preferred strategy in terms of perspective choice (cf. section 2.2.1.5 above).

Altogether, Tversky et al. conclude (1999:409):

> "Ceteris paribus, speakers select reference objects that are salient to communication partners and terms of reference that are relatively easy to produce and comprehend. Switching perspective may enable the more salient reference object and easier reference terms."

These findings may also shed some light on how speakers choose reference frames at all (as addressed in section 2.2.1.4 above, in contrast to perspective selection).

A different approach to identifying speakers' strategies has recently been developed by Fischer (2003), who focusses on the specific situation of linguistic interaction between humans and artificial communicators, i.e., computers or robots. In a setting involving a robot moving to a goal object indicated verbally by a human, Fischer identified several levels of description that represent the users' strategies of achieving their communicative goal. On the highest level, they refer directly to the goal object, using locative dimensional terms. A lower level is to make the robot move towards the goal object by indicating a direction, i.e., by using directionals rather than locative descriptions of the goal object's position. Finally, especially in cases of continuous failure, speakers turn to an even lower level, indicating minor actions like turning on the motor (Fischer & Moratz 2001). This account highlights how speakers' strategies are intricately connected with, on the one hand, the overall discourse situation that they are initially unfamiliar with, and on the other hand, the development of the discourse, involving success or failure with the different kinds of strategy. Furthermore, it shows that it is by no means self-evident that speakers directly refer to a goal object even in a situation where a goal position is to be reached. Rather, the discourse situation may encourage the speaker to turn to more fine-grained, incremental instructions for the same aim.

Speakers' reasons for choosing directional rather than goal-based instructions are as yet not entirely clear. In the present work, several aspects have already been mentioned that might provide some explanation: As Tversky et al. (1999) point out, dimensional terms (used for indicating the spatial position of goal objects) are often hard to process and difficult to employ because of the complexity of factors involved in their interpretation. This fact is mirrored in the present account of the diversity of perspectives and reference systems (as depicted in Figure 6 above) plus the various discoursal and object-related aspects that come into play in the application of locative dimensional terms. These factors become increasingly serious with increasing complexity of the spatial situation, and decreasing referability and discriminability of the objects involved. For example, they may not be situated directly or close to one of the focal axes with respect to a salient relatum, or several competing objects may be situated on the same axis. In contrast, directionals can be used with respect to a moving entity's position

in a fairly uncomplicated and straightforward way, with considerably fewer factors to account for. Thus, it seems reasonable to turn to this simpler strategy especially in cases where uncertainty about the interaction partner's abilities further complicates the situation, as is typically the case in human-robot interaction, at least if the users are not specifically informed about the robot's features.

5.2.2 Figure and Ground

Section 2.6 above showed how, in spatial discourse, entities need to be accessible. The present section addresses the question of how speakers represent accessible entities in discourse, and how they relate them linguistically. Like the previous section, this achievement can be viewed as a way in which speakers employ certain systematic kinds of strategies.

In the present work, I have put forward a systematic account of ways in which spatial entities may be related, starting from the terminology used by Levinson (1996, 2003), and integrating different aspects stemming from different theoretical backgrounds and therefore employing different terminology. One prominent approach that has been mentioned several times already is that advocated, for instance, by Talmy (2000), who uses the terms **Figure** and **Ground** for certain pervasive phenomena in linguistic discourse. In spatial language, these terms correspond quite straightforwardly to Levinson's **referent** and **relatum**. Since Talmy (2000) offers some interesting proposals with regard to how Figure and Ground are chosen in discourse, I will, for the present section, adopt this terminology in order to avoid confusion. Note, however, that similar ideas are also put forward by Langacker (1987, 1991), using the terminology **trajector** (TR) and **landmark** (LM).

Talmy characterises the concepts of Figure and Ground as follows (2000:184):

> "The Figure is a moving or conceptually movable entity whose site, path, or orientation is conceived as a variable the particular value of which is the relevant issue. The Ground is a reference entity, one that has a stationary setting relative to a reference frame, with respect to which the Figure's site, path, or orientation is characterized."

Furthermore, he assigns the following property differences to the two kinds of entities (2000:183):

Table 5. Systematic differences between Figures and Grounds (according to Talmy)

Figure (primary object)	**Ground (secondary object)**
Has unknown spatial (or temporal) properties to be determined	Acts as a reference entity, having known properties that can characterize the primary object's unknowns
More movable	More permanently located
Smaller	Larger
Geometrically simpler (often pointlike) in its treatment	Geometrically more complex in its treatment
More recently on the scene/in awareness	Earlier on the scene/in memory
Of greater concern/relevance	Of lesser concern/relevance
Less immediately perceivable	More immediately perceivable
More salient, once perceived	More backgrounded, once primary object is perceived
More dependent	More independent

The plausibility of the existence of such systematic differences between the Figures and Grounds speakers choose for reference is demonstrated by a simple intuitive example (Talmy 2000:183):

(117) The bike is near the house.

(118) The house is near the bike.

Since the latter example is not likely to be used in natural discourse, while the former seems completely natural, they cannot be treated as synonymous, in spite of the fact that they may both describe the same spatial situation. Herskovits (1986:85) notes that the irreversibility of spatial descriptions is often due to the typicality of the relationships between objects and their use (or functionality), as exemplified by the anomalous example * *the bottle under the cap.*

As far as I know, the details of Talmy's proposals have not been confirmed empirically on the basis of speakers' intuitive choices of Figures and Grounds. Many empirical studies in the area of spatial language dictate which objects are to be related with respect to other objects present in the scenario, by asking about the spatial location of one (specified) object. Exceptions are the studies presented by Taylor et al. (2000, 2001) already described in section 2.2.1.4, who left the decision about choice of referent and relatum to the participants. They found that speakers typically used the more stable object as the referent, and the other object (the less stable one) as the relatum. Relative size did not affect the selection (Taylor et al. 2001).

However, these scenarios are still very much restricted because there were only two objects present. Thus, they are not suitable for identifying speakers' strategies in more open situations. For this aim, corpus-based studies investigating natural occurrences of spatial terms across settings are probably more suited. An example is the recent investigation by de Vega and Rodrigo (2004), who partly support, but partly disprove Talmy's proposal, pointing to some interesting differences between the horizontal and the vertical dimensions. For instance, that Grounds tend to be larger than Figures seems to be true for the vertical dimension but not for the horizontal. In the horizontal dimension, many utterances describe the relative position of two persons to each other (e.g, "John sat behind Mary."). In general, people, animals, and plants are more frequently used for reference in the horizontal than in the vertical dimension. Thus, it seems that the vertical dimension is used for descriptions of different kinds of relations than the horizontal dimension, which may explain the divergence in findings on relative sizes between the dimensions.

Some further (though different) empirical evidence is described by Herrmann & Grabowski (1994), who point out that the choice of Figure and Ground is interrelated with ease of comprehension. For example, a typical strategy is to maintain the same Figure throughout a discourse, yielding a simple thematic structure in which the figure (e.g., a person whose spatial surroundings are described) remains the constant topic and usually syntactic subject (cf. Ehrich's findings discussed in section 2.2.1.4 above).

Another strategy is to rely on a Given / New structure which repeatedly takes the New of the previous sentence as a Given in the next, as in (Herrmann & Grabowski 1994:119):

(119) Am Ende des Flurs ist ein großes Zimmer. In dem Zimmer steht hinten ein Tisch. Auf dem Tisch befindet sich ein Globus. In dem Globus steckt dort, wo Europa ist, eine Stecknadel.

[At the end of the hall there is a big room. In the back of the room there is a table. On the table there is a globe. In the globe there is a pin at the place where Europe is.]

Furthermore, speakers adapt their localisation strategies to the requirements of their listener, in certain situations preferring to use the interlocutor's perspective, and always coordinating their utterances with the listener's reactions. As a consequence of this fact, speakers consider the interlocutor's information status in their utterances as well as their presumed conceptualisations of the spatial situation (Herrmann & Grabowski 1994:117f.):

> "Es ist einleuchtend, daß der Partner den Ort des Relatums schon kennen sollte, bevor das intendierte Objekt durch entsprechende sprachliche Mittel an das Relatum 'angebunden' wird. Wenn der Partner nicht weiß, wo der Stuhl steht, ist es sinnlos zu sagen: 'Der Ball liegt hinter

> dem Stuhl.' (…) Vom Relatum ist in der Regel früher die Rede als vom intendierten Objekt. (…) Welches von zwei Objekten zu[m intendierten Objekt] wird und welches zu[m Relatum], ist zum einen dadurch zu erklären, daß der Sprecher (…) möglichst so formuliert, daß (…) dem Partner das *Verständnis* ermöglicht beziehungsweise erleichtert wird (…) [oder] weil der Sprecher selbst eine Raumkonstellation in bestimmer Weise *auffaßt.*"
>
> [It is evident that the partner should already know the location of the relatum before the intended object is 'bound' to the relatum by suitable linguistic means. If the partner does not know where the chair is, then it is pointless to say: 'The ball is behind the chair.' (…) The relatum is usually mentioned earlier than the intended object. (…) Which of two objects becomes the intended object and which becomes the relatum is, on the one hand, to be explained by the fact that the speaker (…) preferably formulates in a way that (…) the partner's *comprehension* is enabled or made easier (…) [or] because the speakers themselves *conceptualise* the spatial configuration in a specific way.]

Note that these observations are related to – and in accord with – the findings on accessibility discussed in section 2.6 above. Thus, altogether the choice of Figure and Ground in discourse interacts with discoursal strategies as well as conceptualisations of the spatial situation such as those listed by Talmy (2000). For instance, as Herrmann & Grabowski (1994:119) point out, objects viewed as surrounding other objects more often serve as Ground, while surrounded objects usually come to be the Figure. Movable objects are preferably classified as Figure (see also Ehrich 1985:141). If the only difference concerns size then the smaller object is chosen as Figure. Contextually salient objects will also be referred to as Figure.

Further evidence, especially with respect to the aspect of contextual saliency, comes from the literature on the choice of landmarks in route descriptions. Reliably, only those entities are chosen for reference that are salient in the spatial setting and easily identifiable (e.g., Michon & Denis 2001).

5.3 Aspects of the discourse situation

In this section, two aspects of the discourse situation are briefly addressed which have been recognized in the literature to have a strong impact on speakers' linguistic choices. However, since the specific influence on the application of spatial dimensional terms has not been addressed in the literature so far, the following subsections remain fairly abstract, leaving a more concrete investigation to the empirical work currently conducted in our project I1-[OntoSpace].

5.3.1 Adaptation to the interaction partner

Over the past decades there have been increasing efforts in investigating the ways in which speakers adapt to their interaction partner, partly on the grounds of pre-conceptions about the interlocutor and partly – dynamically – via negotiation processes (e.g., the interactive alignment effects investigated by Pickering & Garrod 2004). In this regard, several aspects have already been mentioned in the present work: for instance, in choosing a perspective it may be decisive who the speaker is talking to (cf. section 2.2.1.5 above). Furthermore, the discourse history – a history of successful or unsuccessful communication, of feedback and establishment of common ground – may influence all aspects of linguistic choices, as mentioned, for example, in considering the discourse task of achieving spatial referential identification (section 5.1). These processes also entail the alignment of reference frames to previous usage (Watson et al. 2004, Tenbrink & Moratz 2003). And finally, for each linguistic choice it is decisive whether the speaker considers their interaction partner able to comprehend and interpret the relevant underlying conceptions. Crucially, in a spatial description that is based on a visual perspective, which has been shown to be the case in most usages of dimensional expressions, it is necessary to believe that the interaction partner possesses perceptual abilities in the first place. Such considerations are especially interesting

in a context involving natural interaction between humans and robots, as currently investigated empirically in the project I1-[OntoSpace] (see e.g., Fischer 2004). For example, users' strategy choices are influenced by whether the robot initially presents a scene description to the user, i.e., indicates what it perceives (Moratz & Tenbrink, subm.).

5.3.2 Written vs. spoken mode

Another factor of the discourse situation that is immediately motivated and addressed by the project work in I1-[OntoSpace] is the differentiation of spoken and written mode in human-robot interaction. While most situations considered in the present report concern spatial reference in a typical face-to-face situation involving the spoken mode, human-robot interaction is often carried out by typing instructions into a console. However, the spoken mode may also be employed (involving additional technical problems with consequences of their own), which raises the interesting question of whether similar kinds of language are to be expected in both modes. Thus, the particular – yet at present not extraordinary – case of human-robot dialogues carried out in written mode due to a lack of suitable technology for spoken interaction represents a new kind of discourse situation with relevance to the field of text typology.

In this field, it has been established that texts differ with regard to some generalizable linguistic properties, depending on situational, discourse-related, or content-related factors such as formal vs. informal, spoken vs. written, or monologic vs. dialogic discourse. Much discussion here has concerned the spoken/written distinction (e.g., Biber 1988). There are reasons to believe that this distinction is not per se decisive regarding the presence or absence of linguistic elements, but rather, that there are other factors tending to co-occur in spoken or written texts, that cause certain similarities and differences between text genres that are generally either spoken or written.

Human-robot interaction in written mode, i.e., situations where human users are required to communicate with a robot by typing instructions on a keyboard, can be regarded as a unique situation, characterised by situational factors that are new and unfamiliar to the human users. With human and robot both present, the scenario might be characterised as a face-to-face situation. However, in normal face-to-face situations, the participants communicate using spoken, not written, language. Addressing a present interaction partner using written language is a situation that is only known to the human users from exceptional settings, e.g., when one of the participants is hearing-impaired. Moreover, present participants are usually capable of perceiving their interaction partner and interpreting alterations of facial expression and gestures, as well as reacting to them, often without conscious awareness. Most modern robots are not designed for these aspects of communication, even if they can deal with spoken language.

From the point of view of text typology, such an interaction situation can be regarded as representing an interesting new text type, that is being established under the combined effort of system designers (usually unfamiliar with linguistic insights of this kind), and test users who try to achieve mutual understanding with their robot interaction partner. Previous work in text typology offers several clues regarding the kind of language that might be expected in this kind of text; however, as there has been no systematic analysis with this specific focus, there is insufficient material to be able to make detailed predictions. Viewed this way, human-robot interaction scenarios offer a unique opportunity to watch the development and conventionalisation of a new text type. The crucial point here is that users are uncertain in their communication with a robot (especially in written mode) precisely because they do not know anything about this kind of text, as it did not exist in a standard form previously. Most importantly, the question needs to be addressed whether switching from spoken to written mode is as natural to human users as generally seems to be implicitly assumed. Conceivably,

certain communicative aims are hindered or hampered by the written mode. Certain strategies might not be employed as easily as in oral communication, etc. On the other hand, communication problems will not be attributed to acoustic factors, as is predominant in spoken human-computer interaction. Thus, users will explore different strategies in order to make themselves understood. Since some of these strategies may affect the application of dimensional terms, it is also a specific – though not central – concern of the work in the project I1-[OntoSpace] to identify some effects of choice of mode on the employment of spatial terms. First results, indicating that (surprisingly) the mode of interaction does not fundamentally influence speakers' choices, are presented in Moratz & Tenbrink (2003).

6 Conclusion

In the present work, I have summarized and explored findings on the syntax, semantics, and application (pragmatics) of dimensional terms. While not assuming that these three fields represent distinct and independent areas of investigation, I have worked out some interrelations between syntactic appearances and associated meanings as well as possible applications in the diverse kinds of reference systems that build the basis for the interpretation of the expressions under investigation. I have presented a systematic overview of available reference systems for locative dimensional terms. Furthermore, I have worked out how the lexical concepts can be put to use in diverse discourse contexts dependent on the spatial setting and in interrelation with the objects to be related linguistically.

One aim of this report has been to point to gaps in the existing literature, motivating the empirical research carried out in I1-[OntoSpace] as well as highlighting further aspects that may fall beyond the scope of this work.

Among the issues still underrepresented in the literature so far are at least the following:

- What kinds of factors are crucial for the employment of locative dimensional terms rather than other kinds of options, such as directionals, distance-related, or topological expressions?
- How do speakers choose a reference system in a situation that allows for different kinds of linguistic representations? What kinds of spatial or discourse-related factors influence such decisions?
- How explicit are speakers with regard to underlying reference systems and perspectives in light of the fact that many different factors come into play in the interpretation of dimensional terms?
- What are the effects of a discourse task involving referential identification rather than localisation?
- Under what circumstances does group-based relative reference come into play, and how is it represented linguistically?
- What kinds of factors determine the level of specificity and other aspects related to the notion of granularity in the application of dimensional terms?
- Can the observations on differences between prepositions in English (e.g., ***at*** *the right* vs. ***to*** *the right*) with regard to topology and proximity be confirmed?
- Do speakers combine different conceptualisations of a spatial scene in their spontaneous utterances, and if so, how?
- What kinds of applicability conditions can be identified for directionals?
- What effects does the special situation of spoken or written human-robot interaction have on the application of dimensional terms? For example, given that descriptions of spatial relations depend on the speaker's conceptualisation of the situation, how do speakers react in a situation where the addressee's conceptualisation is unknown?

Acknowledgements

This work is part of the project work of I1-[Onto-Space] of the Transregional Collaborative Research Center SFB/TR 8 Spatial Cognition: Reasoning, Action, Interaction (Bremen / Freiburg, Germany). Funding by the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged. I wish to thank John Bateman and Christian Freksa for very helpful comments on earlier versions of this report, and Reinhard Moratz for numerous inspiring discussions in the area of reference systems.

References

Aurnague, M. and L. Vieu. 1993. A three-level approach to the semantics of space. In C. Zelinsky-Wibbelt (ed.), *The Semantics of Prepositions From Mental Processing to Natural Language Processing,* pp. 393-439. Berlin: Mouton de Gruyter.

Bateman, J. and S. Farrar. 2004. *Spatial Ontology Baseline*. Deliverable D2, I1-[OntoSpace], SFB/TR8 Spatial Cognition, University of Bremen.

Beun, R.-J. and A. Cremers. 2001. Multimodal Reference to Objects: An Empirical Approach. In H. Bunt, R.-J. Beun (eds.), *Cooperative Multimodal Communication*. Berlin, Heidelberg: Springer.

Biber, D. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

Brown-Schmidt, S. and M.K. Tanenhaus. 2003. Referential domains and the interpretation of referring expressions in interactive conversation. *Proc. DiaBruck, 7th Workshop on the Semantics and Pragmatics of Dialogue*, Sept. 4th-6th 2003, Wallerfangen, pp 15-19.

Brugman, C. 1988. *The Story of OVER: Polysemy, Semantics and the Structure of the Lexicon*. New York: Garland Press.

Bryant, D.J., B. Tversky, and M. Lanca. 2000. Retrieving spatial relations from observation and memory. In E. van der Zee and U. Nikanne (eds.), *Cognitive interfaces: Constraints on linking cognitive information.* Oxford: Oxford University Press, 94-115.

Buhl, H.M., S. Katz, K. Schweizer, and T. Herrmann. 2000. Einflüsse des Wissenserwerbs auf die Linearisierung beim Sprechen über räumliche Anordnungen. In: *Zeitschrift für Experimentelle Psychologie, 47*, 1, 17-33. Göttingen: Hogrefe.

Burigo, M. and K. Coventry (2005). Reference Frame Conflict in Assigning Direction to Space. In Freksa, C., M. Knauff, B. Krieg-Brückner, B. Nebel, and T. Barkowsky (eds.), *Spatial Cognition IV: Reasoning, Action, Interaction*. International Conference Spatial Cognition 2004, Frauenchiemsee, Germany, October 2004, Proceedings. Berlin, Heidelberg: Springer, pp. 111-123.

Carlson, L.A. 1999. Selecting a reference frame. *Spatial Cognition and Computation,* 1 (4), 365-379.

Carlson-Radvansky, L.A. and D.E. Irwin. 1993. Frames of reference in vision and language: Where is above? *Cognition 46*, 223-244.

Carlson, L., T. Regier, and E. Covey. 2003. Defining Spatial Relations: Reconciling Axis and Vector Representations. In E. van der Zee and J. Slack (eds.), *Representing Direction in Language and Space*. Oxford: Oxford University Press.

Carlson, L. and E. van der Zee (eds.). 2004. *Functional features in language and space: Insights from perception, categorization and development.* Oxford: Oxford University Press.

Carlson-Radvansky, L.A., E.S. Covey, & K.M. Lattanzi. 1999. "What" effects on "where": Functional influences on spatial relations. *Psychological Science, 10,* 516-521.

Carlson-Radvansky, L.A. and G.D. Logan. 1997. The Influence of Reference Frame Selection on Spatial Template Construction. In: *Journal of Memory and Language* 37, pp. 411-437.

Carlson-Radvansky, L.A. and G.A. Radvansky. 1996. The influence of functional relations on spatial term selection. *Psychological Science, 7,* 56-60.

Carroll, M. 1993. Deictic and intrinsic orientation in spatial descriptions: a comparison between English and German. In J. Altarriba (ed.), *Cognition and Culture: A Cross-Cultural Approach to Psychology.* Elsevier.

Carroll, M. 1997. Changing place in English and German: language-specific preferences in the conceptualization of spatial relations. In J. Nuyts and E. Pederson (eds.), *Language and Conceptualization*. Cambridge University Press, pp. 137-161.

Carston, R. 2002. *Thoughts and Utterances: The pragmatics of explicit communication.* Blackwell.

Clark, H.H. 1973. Space, time, semantics, and the child. In: Moore, Timothy E. (ed). *Cognitive Development and the Acquisition of Language*. N.Y. Academic Press.

Clark, H.H. 1996. *Using Language*. Cambridge, UK: Cambridge University Press.

Clark, H. and Wilkes-Gibbs, D. 1986. Referring as a collaborative process. *Cognition*, 22:1–39.

Clementini, E., P. Di Felice, and D. Hernández. 1997. Qualitative representation of positional information. *Artificial Intelligence*, 95:317--356.

Coventry, K.R., A. Cangelosi, R. Rajapakse, A. Bacon, S. Newstead, D. Joyce, and L.V. Richards. 2004. Spatial Prepositions and Vague Quantifiers: Implementing the Functional Geometric Framework. In C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, and T. Barkowsky (eds.), *Spatial Cognition IV: Reasoning, Action, Interaction*. *International Conference Spatial Cognition 2004, Frauenchiemsee, Germany, October 2004, Proceedings*. Berlin, Heidelberg: Springer, pp. 97-110.

Coventry, K. R. and Garrod, S. C. 2004a. *Saying, seeing and acting: The psychological semantics of spatial prepositions*. Psychology Press: Essays in Cognitive Psychology series.

Coventry, K. R. and Garrod, S. C. 2004b. Spatial prepositions and the functional geometric framework. Towards a classification of extra-geometric influences. In L. A. Carlson & E. van der Zee (eds.), *Functional features in language and space: Insights from perception, categorization and development*. Oxford University Press.

Coventry, K.R., M. Prat-Sala, and Richards, L.V. 2001. The interplay between geometry and function in the comprehension of 'over', 'under', 'above' and 'below'. *Journal of Memory and Language* 44, 376-398.

Crawford, E.L., T. Regier, and J. Huttenlocher. 2000. Linguistic and non-linguistic spatial categorization. *Cognition* 75, pp. 209-235.

Dale, R. and E. Reiter. 1995. Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science,* 18:233-263.

van Deemter, K. and S. Peters (eds). 1996. *Semantic Ambiguity and Underspecification*. CSLI, Stanford, CA.

Duwe, I., K. Kessler and H. Strohner. 2002. Resolving Ambiguous Descriptions through Visual Information. In K.R. Coventry and P. Olivier (eds.), *Spatial Language: Cognitive and Computational Perspectives.* Dordrecht: Kluwer Academic Publishers, pp. 43-67.

Egenhofer, M. and Herring, J. 1991. Categorizing binary topological relationships between regions, lines, and points in geographic databases. In M. Egenhofer, J. Herring, T. Smith, and K. Park (eds.), *A framework for the definition of topological relationships and an algebraic approach to spaial reasoning within this framework,* NCGIA Technical Report 91-7.

Ehrich, V. 1985. Zur Linguistik und Psycholinguistik der sekundären Raumdeixis. In Schweizer, H. (Hrsg.). *Sprache und Raum: Ein Arbeitsbuch für das Lehren von Forschung.* Stuttgart: Metzler, S. 130-161.

Emmorey, K. and S. Casey. 2002. Gesture, Thought, and Spatial Language. In K.R. Coventry & P. Olivier (eds.), *Spatial Language: Cognitive and Computational Aspects*. Dordrecht: Kluwer, pp. 87-101.

Eschenbach, C. 2004. Contextual, Functional, and Geometric Components in the Semantics of Projective Terms. In L. Carlson and E. van der Zee (eds.), *Functional features in language and space: Insights from perception, categorization and development*. Oxford: Oxford University Press.

Eschenbach, C. 1999. Geometric structures of frames of reference and natural language semantics. *Spatial Cognition and Computation,* 1 (4), 329-348.

Eschenbach, C., C. Habel, and A. Leßmöllmann. 2001. Multiple frames of reference in interpreting complex projective terms. In K.R. Coventry and P. Olivier (eds.), *Spatial Language: Cognitive and Computational Perspectives*. Dordrecht: Kluwer, pp. 209-231.

Eschenbach, C. and L. Kulik. 1997. An axiomatic approach to the spatial relations underlying *left-right* and in *front of-behind*. In G. Brewka, C. Habel and B. Nebel (eds.), *KI-97: Advances in Artificial Intelligence* (pp. 207-218). Berlin: Springer.

Evans, V. 2004. *The Structure of Time: Language, meaning and temporal cognition*. Amsterdam/ Philadelphia: John Benjamins.

Fischer, K. 2003. Linguistic Methods for Investigating Concepts in Use. In: Stolz, Th. and Kolbe, K. (eds.): *Methodologie in der Linguistik*. Frankfurt a.M.: Lang.

Fischer, K. and R. Moratz. 2001. From Communicative Strategies to Cognitive Modelling. Workshop Epigenetic Robotics, Lund.

Fischer, K. 2004. Addressee-oriented Blending in Spatial Object Localisation Tasks. In T. Barkowsky, C. Freksa, M. Knauff, B. Krieg-Brückner, and B. Nebel (eds.), *Spatial Cognition 2004: Poster Presentations.* Frauenchiemsee, Germany, October 2004. SFB/TR8 Report No. 002-10/2004, Universität Bremen / Universität Freiburg.

Frank, A.U. 1998. Formal models for cognition – Taxonomy of spatial location description and frames of reference. In C. Freksa, C. Habel and K.F. Wender (eds.), *Spatial Cognition. An Interdisciplinary Approach to Representing and Processing Spatial Knowledge* (pp. 293-312). Berlin: Springer-Verlag.

Franklin, N. and T. Federico. 2002. Organization of Temporal Situations. In K.R. Coventry & P. Olivier (eds.), *Spatial Language: Cognitive and Computational Aspects*. Dordrecht: Kluwer, pp. 103-120.

Freksa, C. 1980. Communication about visual patterns by means of fuzzy characterizations. *XXIInd Intern. Congress of Psychology*, Leipzig, July 1980.

Freksa, C. 1981. *Linguistic pattern characterization and analysis*. Dissertation, University of California, Berkeley.

Freksa, C. 1991. Conceptual neighborhood and ist role in temporal and spatial reasoning. In Singh and Travé-Massuyès (eds.), *Decision Support Systems and Qualitative Reasoning.* Amsterdam: North-Holland, pp. 181-187.

Freksa, C. 1992a. Temporal reasoning based on semi-intervals. *Artificial Intelligence, 54*(1-2), 199-227.

Freksa, C. 1992b. Using Orientation Information for Qualitative Spatial Reasoning. In A. U. Frank, I. Campari, and U. Formentini, editors, *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, pages 162--178. Springer-Verlag, Berlin, Heidelberg, New York.

Freksa, C. 1999. Links vor - Prototyp oder Gebiet? Probabilistische und possibilistische Raumbeschreibung. In G. Rickheit (ed), *Richtungen im Raum*, 231-246, Westdeutscher Verlag, Wiesbaden

Friederici, A. and W.J.M. Levelt. 1990. Spatial reference in weightlessness: Perceptual factors and mental representations. *Perception & Psychophysics* 47(3):253-66.

Gapp, K.P. 1994a. Basic Meanings of Spatial Relations: Computation and Evaluation in 3D Space. Proc. AAAI 1994, Seattle, WA, pp. 1393-1398.

Gapp, K.P. 1994b. A computational model of the basic meanings of graded composite spatial relations in 3d space. *Proc. of the AGDM'94*, Delft, The Netherlands, 66-79.

Gapp, K.-P. 1995. An empirically validated model for computing spatial relations. In I. Wachsmuth, C.R. Rollinger, and W. Brauer (eds.): *KI-95: Advances in Artificial Intelligence. 19th Annual German Conference on Artificial Intelligence* (pp 245-256). Springer: Berlin, Heidelberg.

Garnham, A. 1999. What's in a Mental Model? In G. Rickheit and C. Habel (eds.), *Mental Models in Discourse Processing and Reasoning*. Amsterdam: Elsevier.

Garrod, S. and Anderson, A. 1987. Saying what you mean in dialogue: A study in conceptual and semantic coordination. *Cognition* 27: 181-218.

Garrod, S., G. Ferrier, and S. Campbell. 1999. *In* and *on*: investigating the functional geometry of spatial prepositions. *Cognition* 72 (1999), 167-189.

Garrod, S.C. and A.J. Sanford. 1988. Discourse models as interfaces between language and the spatial world. *Journal of Semantics* 6: 147-160.

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gorniak, P. and D. Roy. 2004. Grounded Semantic Composition for Visual Scenes. *Journal of Artificial Intelligence Research*, Volume 21, pages 429-470.

Grabowski, J. 1998. Ein psychologisch-anthropologisches Modell der einheitlichen semantischen Beschreibung dimensionaler Präpositionen. In P. Ludewig and B. Geurts (eds.), *Lexikalische Semantik aus kognitiver Sicht – Perspektiven im Spannungsfeld linguistischer und psychologischer Modellierungen* (S. 11-40). Tübingen: Narr.

Grabowski, J. 1999. Raumrelationen: Kognitive Auffassung und sprachlicher Ausdruck. Wiesbaden: Westdeutscher Verlag.

Grice, H.P. 1975. Logic and conversation. In: Cole, P. and Morgan, J. (eds.), *Syntax and semantics*. New York, San Francisco, London (Vol 3: 41-58).

Grosz, B.J. and Sidner, C.L. 1986. Attention, intentions and the structure of discourse. *Computational Linguistics* 12(3):175-204.

Gryl, A., B. Moulin and D. Kettani. 2002. A conceptual model for representing verbal expressions used in route descriptions. In Kenny R. Coventry and Patrick Olivier (eds.), *Spatial Language: Cognitive and Computational Perspectives.* Dordrecht: Kluwer Academic Publishers, pp. 19-42.

Habel, C. 1985. Prinzipien der Referentialität. Habilitationsschrift, TU Berlin.

Habel, C. 1999. Drehsinn und Reorientierung – Modus und Richtung beim Bewegungsverb *drehen*. In: G. Rickheit (Hrsg.), *Richtungen im Raum.* Opladen: Westdeutscher Verlag.

Hayward, W.G., and Tarr, M.J. 1995. Spatial language and spatial representation. *Cognition 55,* pp. 39-84.

Hernández, D. 1994. *Qualitative Representation of Spatial Knowledge*, Vol. 804 of *Lecture Notes in Artificial Intelligence*. Springer, Berlin.

Herrmann, T. 1990. Vor, hinter, rechts und links: das 6H-Modell. Psychologische Studien zum sprachlichen Lokalisieren. *Zeitschrift für Literaturwissenschaft und Linguistik 78.* 117-140.

Herrmann, T. und W. Deutsch. 1976. *Psychologie der Objektbenennung*. Bern u.a.: Hans Huber Verlag.

Herrmann, T. und J. Grabowski. 1994. *Sprechen. Psychologie der Sprachproduktion.* Heidelberg: Spektrum.

Herskovits, A. 1986. *Language and spatial cognition.* Cambridge University Press.

Herzog, G. 1995. *Coping with Static and Dynamic Spatial Relations.* In: P. Amsili, M. Borillo, L. Vieu (eds.), Proc. of TSM'95, Time, Space, and Movement: Meaning and Knowledge in the Sensible World, pp. 47-59.

Hill, C.A. 1982. Up/down, front/back, left/right. A contrastive study of hausa and English. In J. Weissenborn and W. Klein (eds.), *Here and There. Cross-linguistic Studies on Deixis and Demonstration* (pp. 13-42). Amsterdam: John Benjamin Publishing Company.

Hobbs, J.R. 1995. Sketch of an ontology underlying the way we talk about the world. *International Journal of Human-Computer Studies 43*, 5/6, 819-830.

Johnston, J. 1984. Acquisition of locative meanings: *Behind* and *in front of. Journal of Child Language* 11:407-22.

Kessler, K., I. Duwe & H. Strohner. 1996. Sprachliche Objektidentifikation in ambigen Situationen: Empirische Befunde. *Report 96/1*, SFB 360, Bielefeld.

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. 1991. Raumausdrücke. *Linguistische Berichte, 132,* 77-114.

Kray, C. 2003. *Situated interaction on spatial topics*. Dissertation, Fachbereich Informatik, Universität Saarbrücken.

Kühnlein, P. and J. Stegmann. 2003. Empirical Issues in Deictic Gesture: Referring to objects in simple identification tasks. *Report 2003/3, SFB 360*, Univ. Bielefeld.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind.* Chicago: Chicago University Press.

Landau, B. 2003. Axes and Direction in Spatial Language and Spatial Cognition. In E. van der Zee and J. Slack (eds.), *Representing Direction in Language and Space*. Oxford: Oxford University Press.

Landau, B. and R. Jackendoff. 1993. „What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences 16,* 217-265.

Landragin, F. and L. Romary. 2003. Referring to Objects through Sub-Contexts in Multimodal Human-Computer Interaction. Proc. *DiaBruck, 7th Workshop on the Semantics and Pragmatics of Dialogue*, Sept. 4th-6th 2003, Wallerfangen, pp 163f.

Lang, E. 1989. The Semantics of Dimensional Designation of Spatial Objects. In: M. Bierwisch and E. Lang (eds.), *Dimensional adjectives: Grammatical Structure and Conceptual Interpretation.* Berlin et al.: Springer.

Langacker, R.W. 1986. An Introduction to Cognitive Grammar. *Cognitive Science 10*, 1-40.

Langacker, R.W. 1987, 1991. Foundations of cognitive grammar. Vol I (1987): Theoretical Prerequisites. Vol II (1991): Descriptive Application. California: Stanford University Press.

Levelt, W.J.M. 1996. Perspective Taking and Ellipsis in Spatial Descriptions. In P. Bloom, M.A. Peterson, L. Nadel and 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 and 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. and W. Labov. 1975. Spatial networks as a site for the study of language and thought. *Language 51,* pp 924-939.

Logan, G.D., and Sadler, D.D. 1996. A computational analysis of the apprehension of spatial relations. In P. Bloom, M.A. Peterson, L. Nadel and M.F. Garrett (eds.), *Language and Space*. Cambridge, MA: MIT Press, pp. 493-529.

Mainwaring, S.D., B. Tversky, M. Ohgishi, and D.J. Schiano. 2003. Descriptions of simple spatial scenes in English and Japanese. *Spatial Cognition and Computation 3(1)*, pp. 3-42.

Manning, C., M.D. Sera, and H.L. Pick, Jr. 2002. Understanding how we think about space. In K.R. Coventry and P. Olivier (eds.), *Spatial Language: Cognitive and Computational Perspectives.* Dordrecht: Kluwer Academic Publishers, pp. 147-164.

Marr, D. 1982. *Vision: a computational investigation into the human representation and processing of visual information.* New York: Freeman.

McIntyre, A. 2001. Functional interpretations: Borderline idiomaticity in prepositional phrases, complex verbs and other expressions. Ms.

Michon, P.E. and M. Denis. 2001. When and why are visual landmarks used in giving directions? In D.R. Montello (ed.): *Spatial Information Theory: Foundations of Geographic Information Science*. International Conference, COSIT 2001 Morro Bay, CA, USA, September 19-23, 2001. Proceedings. Berlin, Heidelberg: Springer.

Moratz, R. and T. Tenbrink. 2003. Instruction modes for joint spatial reference between naive users and a mobile robot. *Proceedings of the IEEE International Conference on Robotics, Intelligent Systems and Signal Processing (RISSP 2003), Special Session on New Methods in Human Robot Interaction*, October 8-13, 2003, Changsha, Hunan, China.

Moratz, R. and T. Tenbrink (subm.). Spatial reference in linguistic human-robot interaction: Iterative, empirically supported development of a model of projective relations. Under review.

Moratz, R. and K. Fischer. 2000. Cognitively Adequate Modelling of Spatial Reference in Human-Robot Interaction. *Proceedings of the 12th IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2000*, November 13-15 2000, Vancouver, British Columbia, Canada, pp. 222-228.

Mukerjee, A. 1998. Neat Versus Scruffy: A Review of Computational Models for Spatial Expressions. In Olivier, Patrick and Klaus-Peter Gapp (eds.) *Representation and Processing of Spatial Expressions*. Mahwah, New Jersey: Lawrence Erlbaum, pp. 1-35.

Nikanne, U. 2003. How Finnish postpositions see the axis system. In E. van der Zee and J. Slack (eds.), *Representing Direction in Language and Space*. Oxford: Oxford University Press, pp. 191-208.

Oviatt, S. 1999. Ten myths of multimodal interaction. *Communications of the ACM*, 42(11):74 - 81.

Pederson, E. 2003. How many reference frames? In C. Freksa, W. Brauer, C. Habel, and K.F. Wender (eds): *Spatial Cognition III: Routes and Navigation, Human Memory and Learning, Spatial Representation and Spatial Learning.* Berlin: Springer, pp. 287-304.

Piaget, J. and Inhelder, B. 1956. *The child's conception of space.* London: Routledge and Kegan Paul.

Pickering, M. and Garrod, S. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27.*

Pobel, R., Grosser, Ch., Mangold-Allwinn, R., und Herrmann, Th. (1988). Zum Einfluß hörerseitiger Wahrnehmungsbedingungen auf die Überspezifikation von Objektbenennungen. *Arbeiten der Forschergruppe "Sprechen und Sprachverstehen im sozialen Kontext" Heidelberg/Mannheim, Bericht Nr. 17.*

Porzel, R., M. Jansche and R. Meyer-Klabunde. 2002. Generating spatial descriptions from a cognitive point of view. In K.R. Coventry and P. Olivier (eds.), *Spatial Language: Cognitive and Computational Perspectives.* Dordrecht: Kluwer Academic Publishers, pp. 185-207.

Pribbenow, S. 1992. *Zur Verarbeitung von Lokalisierungsausdrücken in einem hybriden System*. Dissertation. IWBS-Report 211, April 1992, IBM.

Randell, D., Cui, Z. and Cohn, A. 1992. A spatial logic based on regions and connection. In: *Proceedings of the 3rd International Conference on Knowledge Representation and Reasoning,* Morgan Kaufmann, San Mateo, pp. 165-176.

Regier, T. and L. Carlson. 2001. Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General*, 130(2), 273-98.

Regier, T., L. Carlson, and B. Corrigan. 2004. Attention in spatial language: Bridging geometry and function. In Carlson, L. and E. van der Zee (eds.), *Functional features in language and space: Insights from perception, categorization and development*. Oxford: Oxford University Press.

Retz-Schmidt, G. 1988. Various views on spatial prepositions. *AI Magazine 9*: 2. 95-105.

Rieser, H. 1996. Repräsentations-Metonymie, Perspektive und Koordination in aufgabenorientierten Dialogen. In: Carla Umbach, Michael Grabski und Robin Hörnig (ed), *Perspektive in Sprache und Raum.* Wiesbaden: Deutscher Universitäts-Verlag.

Rieser, H. 2004. Pointing in dialogue. *Proceedings of Catalog 2004: Semantics and Pragmatics of Dialog, Barcelona, July 19-21, 2004.*

Rosch, E. 1978. Principles of categorization. In: E. Rosch and B.B. Lloyd (eds.), *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum, pp. 27-48.

Roy, D. 2002. Learning Words and Syntax for a Visual Description Task. *Computer Speech and Language*, 16(3).

Schmidtke, H., L. Tschander, C. Eschenbach, and C. Habel. 2003. Change of orientation. In Emile van der Zee & Jon Slack (eds.), *Representing Direction in Language and Space*. Oxford: Oxford University Press, pp. 166-190.

Schober, M.F. 1998. How addressees affect spatial perspective choice in dialogue. In P.L. Olivier and K.-P. Gapp (eds.), *Representation and processing of spatial expressions* (pp. 231-245). Mahwah, NJ: Lawrence Erlbaum.

Schober, M.F. 1993. Spatial Perspective-Taking in Conversation, *Cognition* 47: 1-24.

Shariff, A.R., M. Egenhofer, and D. Mark. 1998. Natural-language spatial relations between linear and areal objects: The topology and metric of English-language terms. *International Journal of Geographical Information Science 12(3)*, 215-246.

Siegel, A.W., and White, S.H. 1975. The development of spatial representations of large-scale environments. In H.W. Reese (ed.), *Advances in child development and behavior*. New York: Academic Press, pp. 9-55.

van der Sluis, I. and E. Krahmer. Generating Referring Expressions in a Multimodal Context: An empirically motivated approach. In: W. Daelemans et al. (eds.), *Computational Linguistics in the Netherlands 2000, Selected Papers from the 11th CLIN Meeting.* Rodopi, Amsterdam.

Sowa, T. and I. Wachsmuth. 2002. Interpretation of shape-related iconic gestures in virtual environments. In: Wachsmuth, I. and T. Sowa (eds): *Gesture and sign language in human-computer interaction.* International Gesture Workshop, GW 2001, London, UK, April 18-20, 2001 Proceedings . Berlin etc.: Springer, pp. 21-33.

von Stutterheim, C., R. Mangold-Allwinn, S. Barattelli, U. Kohlmann, and H. Kölbing. 1993. Reference to objects in text production. In: J. Nuyts and E. Pederson (eds.), *Perspectives on Language and Conceptualization*. Belgian Journal of Linguistics 8, pp. 99-125.

Talmy, L. 1983. How language structures space. In H. Pick and L. Acredolo (eds.), *Spatial orientation: Theory, research, and application.* New York: Plenum.

Talmy, L. 2000. *Toward a Cognitive Semantics*, 2 vols. Cambridge, MA: MIT Press.

Taylor, H.A., S.J. Naylor, R.R. Faust, and P.J. Holcomb. 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. and S.J. Naylor. 2002. Goal-directed effects on processing a spatial environment. In K.R. Coventry and P. Olivier (eds.), *Spatial Language: Cognitive and Computational Perspectives.* Dordrecht: Kluwer Academic Publishers, pp. 233-253.

Taylor, H.A. and Tversky, B. 1992. Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language 31:* 261-292.

Taylor, H. A. and Tversky, B. 1996. Perspective in spatial descriptions. *Journal of Memory and Language, 35,* 371-391.

Taylor, T. E., Gagné, C. L., & Eagleson, R. 2000. Cognitive constraints in spatial reasoning: Reference frame and reference object selection. *American Association for Artificial Intelligence Technical Report SS-00-04* (pp. 168-172), Menlo Park, CA: AAAI Press.

Taylor, T. E., Gagné, C. L. & Eagleson, R. 2001. Producing spatial descriptions: Effects of object familiarity. *Proceedings of the 1st International Symposium on Smart Graphics*, Hawthorne, NY, 69-75.

Tenbrink, T. (in prep.). *Localising objects and events: Discoursal applicability conditions for spatiotemporal expressions in English and German*. Dissertation project, University of Bremen.

Tenbrink, T. 2005. Identifying objects on the basis of spatial contrast: An empirical study. In Freksa, C., M. Knauff, B. Krieg-Brückner, B. Nebel, and T. Barkowsky (eds.), *Spatial Cognition IV: Reasoning, Action, Interaction*. International Conference Spatial Cognition 2004, Frauenchiemsee, Germany, Proceedings. Berlin, Heidelberg: Springer, pp. 124-146.

Tenbrink, T. and R. Moratz. 2003. Group-based Spatial Reference in Linguistic Human-Robot Interaction. *Proceedings of EuroCogSci 2003: The European Cognitive Science Conference*, September 10-13, Osnabrück, Germany, pp 325-330.

Timpf, S., G.S. Volta, D.W. Pollock, and M.J. Egenhofer. 1992. A Conceptual Model of Wayfinding Using Multiple Levels of Abstraction. In I. Campari, A. U. Frank, and U. Formentini (eds.), *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space.* Berlin, Heidelberg: Springer, pp. 348 – 367.

Thorisson, K. 1994. Simulated perceptual grouping: An application to human computer interaction. *Proceedings of the 16th Annual Conference of Cognitive Science Society*, pp. 876-881, Atlanta GA.

Tucker, M. and Ellis, R. 1998. On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology*: Human Perception and Performance, 24,3, 830-846.

Tversky, B. 1999. Spatial Perspective in Descriptions. In P. Bloom, M.A. Peterson, L. Nadel and M.F. Garrett (eds.), *Language and Space* (pp. 109-169). Cambridge, MA: MIT Press.

Tversky, B. and P.U. Lee. 1998. How Space Structures Language. In: C. Freksa, C. Habel & K.F. Wender (eds.), *Spatial Cognition. An Interdisciplinary Approach to Representing and Processing Spatial Knowledge* (pp. 157-175).

Tversky, B., P. Lee, and S. Mainwaring. 1999. Why do speakers mix perspectives? *Spatial Cognition and Computation*, 1 (4), 399-412.

Tyler, A. and V. Evans. 2003. *The Semantics of English Prepositions: Spatial Sciences, Embodied Meaning, and Cognition.* Cambridge University Press.

de Vega, M. and M.J. Rodrigo. 2004. The bicycle pedal is in front of the table. Why some objects do not fit into some spatial relations. In L. Carlson and E. van der Zee (eds), *Functional and Spatial Features in Language and Space*. Oxford University Press.

Vorwerg, C. 2001. *Raumrelationen in Wahrnehmung und Sprache: Kategorisierungsprozesse bei der Benennung visueller Richtungsrelationen.* Wiesbaden: DUV.

Vorwerg, C. and G. Rickheit. 1998. Typicality effects in the categorization of spatial relations. In: C. Freksa, C. Habel and K.F. Wender (eds.), *Spatial Cognition. An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*. Berlin: Springer-Verlag, pp. 203-222.

Vorwerg, C., G. Socher, T. Fuhr, G. Sagerer, and G. Rickheit. 1997. Projective relations for 3D space: Computational model, application, and psychological evaluation. In: *Proceedings of AAAI 97*, Juli 27-31, 1997, Providence, Rhode Island (pp 159-164).

Vorwerg, C. and Weiß, P. 2004. Einfluss der Verbsemantik auf die Interpretation von Richtungspräpositionen. In: D. Kerzel et al. (Hrsg.), *Experimentelle Psychologie. Abstracts der 46. TeaP in Gießen*, April 2004. Lengerich: Pabst Science Publishers.

Wachsmuth, I. and Y. Cao. 1995. Interactive Graphics Design with Situated Agents. In W. Strasser and F. Wahl (eds.), *Graphics and Robotics*. Berlin etc.: Springer, pp. 73- 85.

Watson, M.E., M.J. Pickering, and H.P. Branigan. 2004. Alignment of Reference Frames in Dialogue. *Cogsci 2004: 26th Annual Meeting of the Cognitive Science Society*, August 5-7, 2004, Chicago.

Weiß, P., J. Grabowski, and G. A. Miller. 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.

von Wolff, A. 2001. *Transformation und Inspektion mentaler Umraumrepräsentationen: Modell und Empirie.* Vienna: GeoInfo Series.

Wunderlich, D. and M. Herweg. 1991. Lokale und Direktionale. In: A. von Stechow and D. Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*. Berlin/New York: de Gruyter.

Wünstel, M. and R. Moratz. Automatic object recognition within an office environment. In *Canadian Conference on Computer and Robot Vision (CRV2004)*, 2004.

van der Zee, E. and R. Eshuis. 2003. Directions from Shape: How Spatial Features Determine Reference Axis Categorization. In E. van der Zee and J. Slack (eds.), *Representing Direction in Language and Space*. Oxford: Oxford University Press, pp. 209-225.

van der Zee, E. and J. Slack (eds.), 2003. *Representing Direction in Language and Space*. Oxford: Oxford University Press.

Zimmer, H.D., H. Speiser, and J. Baus. 2001. Die Selektion dimensionaler räumlicher Präpositionen: automatisch und nicht ressourcenadaptierend. *Kognitionswissenschaft* 9 (3), pp 114-121.

Zimmer, H.D., H.R. Speiser, J. Baus, A. Blocher, and E. Stopp. 1998. The Use of Locative Expressions in Dependence of the Spatial Relation between Target and Reference Object in Two-Dimensional Layouts. In C. Freksa, C. Habel and K.F. Wender (eds.), *Spatial Cognition. An Interdisciplinary Approach to Representing and Processing Spatial Knowledge* (pp. 223-240). Berlin: Springer-Verlag.

Zwaan, R.A. and G.A. Radvansky. 1998. Situation Models in Language Comprehension and Memory. *Psychological Bulletin 123*, 2, 162-185.

Zwarts, J. 1995. The Semantics of Relative Position. *Proceedings of Semantics and Linguistic Theory*, SALT5, 405-422.