[Paper presented at the 3rd ANNUAL LANGUAGE & SPACE WORKSHOP 'Resolution and Granularity in Language and Space'. ZIF, Univ. Bielefeld. 8./9.7.2002]

Location and Distance in Language: An attention-based approach

Kai-Uwe Carstensen

Institute of Computational Linguistics

University of Zurich, Switzerland

carstens [at] ifi.unizh.ch

1. Introduction

Location and Distance are central concepts in the investigation of spatial language and cognition. Location and distance information expressed in language allows a hearer to *localize* an object (i.e., to narrow down the search domain of finding the place of that object). For example, we use locative prepositions for the description of an object´s location (*X is left of Y*), and adjectives and measure phrases for the description of distances (*X is high above Y, X is 10m behind Y*, *X close to Y*).

Two important aspects are involved in the investigation of location and distance in language(see also Peterson et al. 1996, p. 572). One concerns the actual properties of spatial relations and spatial language in use. In this case, the focus of interest is put on the extensional (or „quantitative“) range of spatial expressions (e.g., the size and form of spatial regions, see Carlson, this volume) or, for example, on vector map realizations of spatial relation knowledge in the mind/brain (O´Keefe, 1996). The other aspect is the semantic (or „qualitative“) content of spatial expressions, where the main interest (as for example in this paper) lies in the construction of (cognitive) semantic theories and models that best meet the linguistic phenomena observed.

There is a widely accepted “classical” view of (linguistic) spatial relations, in which linguistic phenomena of localization are modelled by region inclusion. Yet some relevant phenomena are not adequately handled in this view, one of them being the so-called „modification problem“: the observation that there are restrictions to the combination of location and distance in language which cannot be explained on purely extensional grounds. In the current paper, an alternative to this view is presented.

Parallel to the growing work in formal spatial semantics, a fruitful line of reasearch has been established (e.g., Miller & Johnson-Laird, 1976, Olson & Bialystok, 1983, Jackendoff, 1983, Landau & Jackendoff, 1993, Bloom et. al., 1996) which is based on the assumption that it is useful and even necessary to investigate *spatial cognition* for the explanation of spatial semantic phenomena, and that the analysis of the latter is at the same time constructive as a window to the former (cf., e.g., Lang & Carstensen & Simmons, 1991). The present paper is in the tradition of these approaches. It focusses on the question of which aspects of (spatial) cognition are relevant for the specification of spatial semantic representations such that existing subclassifications of spatial prepositions can be cognitively motivated and specific phenomena in the compatibility of distance and location phrases can be explained. Classical approaches do not answer this question satisfactorily. According to the position taken here this is due to a missing level of representation and processing in their theories, namely the working of selective spatial attention.

It has already been shown that attention is necessary for the apprehension of spatial relations (Logan, 1995). However, its role as an interface between visuo-spatial (depictional) and conceptual (propositional) representations has been less considered so far. In contrast to that, I propose that an understanding of how we attend to spatial entities is crucial for modelling (linguistic) spatial relations. According to the view advocated here, they derive both from perceiving and attentionally perspectivizing space. Attentional perspectivization (which will be called "microperspectivization") leads to the establishment of so-called "explicit" spatial relations (as opposed to "implicit" relations in a spatial medium) which are categorized according to a set of qualitative features. It will be argued that existing approaches can be generalized by taking into account qualitative aspects

of microperspectivization deriving from so-called “mental presentation”, which leads to a different proposal for the semantics of spatial expressions.

2. Linguistic Phenomena

There is a widely accepted subclassification of locative prepositions (i.e., those that answer the question *Where is X?*) into *topological* (e.g., English the prepositions *on*, *at*, *by*, *near*, German *an*, *bei*) and *projective* (*above*, *below*, *left of* etc.) ones. Most of the research in this domain seems to have been concentrated on projective prepositions. A few notions (like, e.g., *reference frame*, *spatial axis* and *region (of acceptability)*) have been identified as criterial, and (cross-linguistic) differences in the semantics and use of prepositions can be explained by the variations of how they are imposed on a given reference object, a viewer, or the environment (Herskovits, 1986, Levinson, 1996). Various approaches for modelling these aspects have been proposed, e.g., logic-based (Aurnague & Vieu, 1993), qualitative/propositional (Hernández, 1994), depictional (Glasgow & Papadiaz, 1992), hybrid depictional and propositional (Latecki & Pribbenow, 1992) and connectionist (Regier, 1995) ones.

The situation is quite different for the semantics of topological prepositions, which is notoriously difficult to specify in strictly spatial terms. Because of that, conceptual or *functional* notions (e.g., CONTAINMENT, CONTACT, SUPPORT) have been introduced in the semantics of topological prepositions (Talmy, 1983, Herkovits, 1986, Aurnague & View, 1993, Coventry et al., 1994, Garrod et al., 1999).

Compared with location expressions, distance expressions have received only scarce interest in the investigation of language and space. This may be due to the seemingly simple linguistic structure for coding distance information (adjectives like *near* and *far*, and measure phrases as in *10m above*). A closer look at the linguistic data, however, reveals a more complicated picture. First, distance information is also expressed by prepositions, and it is these spatial terms that pose the most notorious problems in spatial semantics (i.e., the distinction between *at*, *by*, *near*, or German *an*, *bei*). Second, expressions like *far from* and *close to* are often

treated as prepositions. Yet this runs counter to the observation that they consist of a distance adjective modifying a prepositional construction (compare *far above*, *close by*). But then, what about **The restaurant is from the church* or **The restaurant is to the church*, where stripping off the distance modifier does not result in an acceptable location expression? Similar problems exist with measure phrases: Although *100m above the hill* is fine, a question asking for the distance information must start with *How far above* and not with **How above*. Likewise, it is *2 kilometers further down the road* and *How much further?* but not **How further?*. This shows that the mapping from linguistic to conceptual structure and vice versa is not as simple as it may seem. Third, distance information expressed in language could be expected to be only an "add-on" to verbalized location information (as it only presents more details about the spatial relation involved) so that distance and location expression should be quite independent of one another. As has been observed in Carstensen (1992, 1995, 2001), Zwarts (1997), Zwarts & Winter (2000), this is not the case.

(1) a. the bird is 10 meters above/below/beside the tree

b. *the bird is 2 meters near/on the tree

(1) shows that a measure phrase is not compatible with *near*/*on* although distance and location information do not contradict each other. In general, topological prepositions may not be modified by a positive distance expression (**2 meters near*), and modification of a projective preposition with a negative distance expression (?*close above*) is at least questionable. Correspondingly, this compatibility phenomenon has been called *the modification problem in spatial semantics* which represents the main evidence against the classical view of modelling localization in language.

3. The classical view of localization

3.1. General description

Miller and Johnson-Laird state what can be regarded as the main idea underlying the classical view of localization: 'One purpose of locative descriptions is to narrow the domain of search for a referent' (Miller/Johnson-Laird 1976, p. 384). They distinguish two search domains: one to search for the relatum (reference object, RO), and another (sub-)domain to search for the referent (located object, LO) for which '[i]t is necessary , therefore, to specify the subdomain as a function of the preposition and the relatum' (ibid.). For some of the prepositions they analyze, they make use of a concept of REGION of a relatum which represents the region of (conventional) interaction with this object.

Most semanticists have since then used this functional notion of region: According to those (formal) approaches to spatial semantics, spatial prepositional phrases (PPs) either denote a region (or a PLACE, in Jackendoffs terms) or express a property of being located within a region (with different prepositions specifying different regions). In the more common latter case, the semantics of, e.g., *above* would be specified as the two-place predicate in (2) saying 'x is located in the ABOVE*-region with respect to the reference object y' or 'x´s place is part of the ABOVE*-region of y'. This predicate would be interpreted in terms of set theory with an underlying ontology where regions are treated as sets of spatial points, and where the LOC-relation is the *subset*-relation.

(2) above: λy λx [LOC(x, ABOVE*(y))]

The semantics of distance expressions is specified similarly in the classical view, namely as a property of the referent. The combination of location and distance expressions is then treated according to a straightforward compositional pattern: Given a location expression semantically represented as (3a) which is modified by a

distance expression semantically represented as property (3b), the external variables are unified and the predicates are conjoined as in (3c).

(3) a. λx [SOME_LOCATION_PREDICATE(x)]

b. λy [SOME_DISTANCE_PREDICATE(y)]

c. λx [SOME_LOCATION_PREDICATE(x) & SOME_DISTANCE_PREDICATE(x)]

Most researchers in the language and space community agree that the meaning of spatial expressions cannot be described in terms of strictly spatial (geometric and/or topological) criteria only. Therefore, further conceptual or functional conditions like the relation of containment in (4a) or the relation of support in (4b) are assumed to be relevant, especially for topological prepositions.

(4) a. in: λy λx [LOC(x, IN*(y)) & contains(y, x)]

b. on: λy λx [LOC(x, ON*(y)) & supports(y, x)]

3.2. Characteristics of the classical view

There are a few aspects which are characteristic of the classical view and highly relevant for the later argumentation.

Unclear status of 'region'. In the work of Miller and Johnson-Laird, 'region' had a straightforward psychological, explanatory meaning/interpretation (both perceptual and conceptual). Over the years, this seems to have changed to a more linguistically oriented, descriptive interpretation as a 'region of acceptability' of a spatial predicate: 'Roughly speaking, there are three main regions of acceptability: one reflecting good examples,

one reflecting examples that are less than good but nevertheless acceptable, and one reflecting unacceptable ones' (Logan & Sadler, 1996, p. 497). In this sense, regions have obviously lost their explanatory content, only describing the extension of a spatial predicate (i.e., a set of possible places). Thus, the problem of characterizing the semantics of a preposition has simply been shifted to the problem of specifying (acceptable) regions.

In functional accounts of spatial expressions (e.g., Coventry et al., 1994, Garrod et al., 1999), it is shown that locative prepositions sometimes *can* be used in situations where the geometric criteria are not satisfied, and that they sometimes *must not* be used although the criteria are satisfied. Although usually a hybrid account is proposed, the necessity of regions can evidently be questioned.

If less prototypical spatial (or even non-spatial) uses of a preposition are taken into consideration, the unclear status of regions becomes even more apparent. Consider, for instance, the phrase *the knot in the rope*. To speak of a region in which the knot is located here simply begs the question of what this region is (the same is true for a corresponding explanation based on functional containment or, correspondingly, an image-schematic presentation).

According to Landau/Jackendoff, however, regions are necessary constituents of spatial representations: '[...] there is nothing in the image (that we can think of, anyway) that would correspond to the spatial notion of *region* (or, in fact, to the principal axis of a nonrotating sphere). But each of these – whether or not it is mappable from the image – must be encoded in a spatial representation.' (Landau & Jackendoff, 1993, p. 257). This may therefore not leave us with the question *whether* regions are needed in theories of language and space, but, at least, *in which sense*.

Center role of the reference object (RO). Closely related to the region-based conception of the classical view is its assumption that the reference object holds the *center* role in the characterization of linguistic and non-linguistic spatial relations. This center role is determined by the functional notion of region and by the view that

localization is containment of the referent in a specified region. Although selfevident and hardly questionable on first sight, this point is pivotal for the alternative conception of localization presented below.

Problems with distance coding in location expressions. There are topological prepositions which can only be differentiated by means of the typical distance expressed (or, correspondingly, by the size of their regions). This is most clearly exemplified by the German prepositions *an* and *bei* which denote a smaller and bigger region around the centered reference object, respectively (see figure 1). There is a long and ongoing debate on how to capture this distinction (e.g., Herweg, 1991, Li, 1994). Yet it does not seem possible to arrive at a satisfactory result within the classical view.

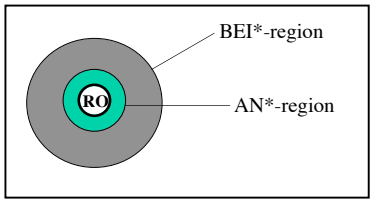


Figure 1: Regions of *an* and *bei*

The modification problem. As has been already mentioned, combinations of distance and location expressions are not adequately handled by classical approaches. For instance, they do not predict the inacceptability of **10 cm an der Wand* (**10 cm close to the wall*). This is due to the fact that the coupling of distance and location expression is too loose in classical semantic composition, for which the restricted ontological base must be blamed. This is stated quite appropriately in a conclusion made by Zwarts/Winter: 'A general compositional treatment of PP modifications is not forthcoming if locative prepositions are taken as relations between sets of points' (Zwarts & Winter, 2000, p. 173).

No resort in functional relations. Sometimes it is argued that problems in the classical view could be solved by resorting to functional relations. For instance, the *an*/*bei*-distinction could be based on only one region and the requirement of *contact* between referent and relatum for *an*. Apart from the fact that this does not solve the modification problem, functional relations even introduce new problems. For example, what does CONTACT mean in the absence of physical contact; what is left of SUPPORT in *stains on the windowpane*; what does CONTAINMENT mean in *knot in the rope*? Rather than trying to adjust the interpretation of functional notions, it may be more adequate to find other, abstract, aspects characterizing (topological) prepositions.

Disregard of implicit/explicit-dichotomy. For quite a while distinctions like propositional/depictional and what/where have predominated cognitive scientific discussions on spatial representations. And yet, none of them has helped much in elucidating the relation of space and language. Only recently the importance of distinguishing implicit and explict spatial relations has been acknowledged:

> ‚Although the [what] system cannot represent explicit spatial relations, it must be able to represent implicit spatial relations; such relations are inherent in any pattern. For example, a representation of a face by necessity includes implicit information about the locations of the parts and the distances among them. However, such spatial representations are embedded in the pattern itself; they cannot be used in any other context' (Kosslyn, 1994, p. 421).

Thus, although the what system contains spatial relations, these relations cannot be used for language. In other words, only explicit spatial relations are relevant for the investigation of language and space. As the (descriptions of) spatial relations in the classical view are mostly concerned with (reference to) space itself, the dichotomy is disregarded there at best and wrongly focused on implicit relations at worst.

4. Vector space semantics

A different approach to the semantics of spatial expressions –more compatible to the one that will be presented here– is given by Joost Zwarts and Yoad Winter. Their modeltheoretic analysis, both of locative prepositions

(Zwarts & Winter, 2000) and of distance expressions (Winter, 2001), is based on spatial *vectors* instead of spatial points (i.e., space ontology consists of a *vector space* V which replaces (or is identified with) the set of spatial points). For their semantics of spatial expressions (called *vector space semantics (VSS)*), the domain D_v of *located vectors* is defined as the cartesian product V×V. Modification of a preposition (to be more exact: of its projection P') can then be straightforwardly modelled as the intersection of the set of located vectors denoted by the P' with the set of located vectors denoted by the modifying distance expression, e.g., *10m above the house* is analyzed as **10m' ∩ above_the_house'**.

In order to cope with the modification problem, the authors define a property of prepositional vectors called *vector monotonicity* which denotes closedness of a vector set wrt. lengthening/shortening (see (5), where ‚≤' is a partial order on D_v).

(5) **Vector monotonicity**

A set of vectors $A \subseteq V$ is *upward (downward) vector-monotone* iff for all vectors $v \in A$ and $w \in V$, if v ≤ w (w ≤ v) then w is in A.

According to this definition, prepositions like *behind/ in front of/ above/ below* can be categorized as being upward vector-monotone and downward vector-monotone, while prepositions like *near/ in/ on/ at/ between* can be identified as being only downward vector-monotone. This corresponds to the fact that distance is irrelevant for the former prepositions, while it evidently matters for the latter.

As to modification of a P', Zwarts & Winter propose that it is acceptable only when the corresponding vector set is upward vector-monotone. They define a modification condition (MC) which states that a set W of located vectors satisfies MC iff for every non-empty set of vectors denoted by an MP (called *measure set*), the intersection M ∩ W is also non-empty. It then follows that W satifies MC iff W is upward vector-monotone, downward vector-monotone, and non-empty.

The analysis, therefore, works quite well for modification of P's by measure phrases. However, this type of modification must be regarded as a special case for the following reasons. First, the modifying structure *can* be an adjective phrase (AP) headed by a distance adjective. In German, for example, it is rather usual to use a distance adjective with the MP (*10m weit unter/über/neben (1m far above/below/beside), einige km hoch über (a few km high above)*). Second, the modifying structure *must* be an AP in questions asking for the description of a comparative distance extent (*How much farther?* vs. **How farther?*) while it *must not* be an AP in the positive case (**How much far?* vs. *How far?*). This shows that the modifying structure of a locative P' should generally be treated as a (if only covert) AP (a view held by, e.g., Bierwisch, 1988, and Carstensen, 1998). Although the modification approach is extended to adjectives in Winter (2001), it remains unclear whether these subleties of spatial semantics are handled adequately in VSS.

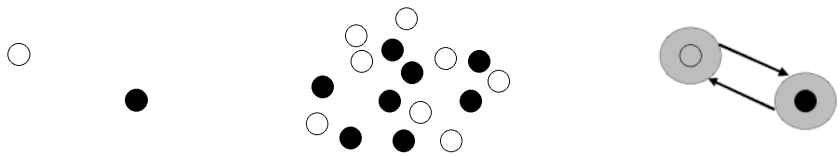
A more serious problem of their approach is the fact that VSS only deals with one part of the modification problem. While it comes out nicely that downward vector-monotone expressions cannot be modified by upward vector-monotone ones, the question remains why it is not possible or at least awkward to use a downward vector-monotone modifier with an upward/downward vector-monotone head (e.g., **very close away/above/behind*). Let us assume that MC is responsible here, too. This would correctly rule out inacceptable expressions but it would also prohibit *any* non-MP description of a short dimensional or distance extent. Yet while this is in fact true for distance extents and some dimensional extents (**close away, *low above, *low man* etc.), there are still *short lines*, *thin boards* etc.

On the whole, VSS represents a big step away from the classical view but still shows some of its characteristics. For instance, it is clearly a RO-centred approach: 'Each "point" *w* in D_p (= a vector in V) functions as ‚the center' (=the zero vector) of a vector space $V_w \subseteq D_v$' (Zwarts & Winter 2000, p. 174). In addition to that, it remains unclear about the status of the formalized relations wrt. the implicit/explicit-dichotomy. It therefore is, like many formal semantic theories, non-cognitive and merely descriptive (a pleonasm

in itself) and so does not contribute much to the discussion of how language and space are linked by conceptual representations.

5. Establishing explicit spatial relations: Microperspectivization

What are explicit spatial relations and how do they get established? Consider the two objects in figure 2a. Obviously, there is an implicit spatial relation between them. The same objects and their relation are also present in figure 2b. However, as is demonstrated clearly, we do not see this relation as easily as in figure 2a. Mere presence of the implicit relation therefore does not imply its availability to the observer (or, correspondingly, to the cognitive processes underlying the recognition of a certain spatial relation). This becomes evident most dramatically in cases of so-called object-based neglect, where patients are not able to see even the spatial relation in figure 2a (although their visual areas are intact). They would report only the existence of one object as their attention is "stuck" to it (Behrmann & Tipper, 1994).



a. b. c.

Figure 2: Spatial Relations and Microperspectives

Regarding the general role of (selective) attention, it has been shown (cf. Theeuwes, 1993) that there is a serial stage at which objects in the visuo-spatial medium (the „visual buffer“ of Kosslyn, 1994) are attentively selected one after another for further processing in the what-system. Exactly what gets attended at a certain point of time is jointly determined by the properties of the given and preprocessed entities of the „display“ (→ bottom up

aspect) and stored patterns of attentional behaviour (attentional templates, → top-down aspect). As to the bottom-up aspect, it is controlled by two main principles: Differences in the display attract the attentional window and thus determine salient entities (bounded regions or *boundaries* of regions) to be further processed, and inhibition of visited places/objects prevents immediate return to those entities.

Shifts of selective attention are therefore necessary for establishing explicit spatial relations ('Computing relations requires directing attention', Logan, 1995, p. 163). This is important, because it marks a characteristic difference to approaches within cognitive linguistics based solely on (only implicitly spatial) image schemata. Another aspect of attention-based relations is shown in figure 2c. As there are two possibilities of attention shifts between the objects of an implicit relation, the shifts have the effect of imposing a certain perspective (which I call „microperspective") on it. {1} These microperspectives themselves – the displacements of the attentional „window" (cf. Kosslyn, 1994) – constitute the core of explicit spatial relations that get verbalized as prepositions. According to this view, then, spatial prepositions express qualitative information about the microperspectivization of spatial representations available in working memory, rather than implicit spatial information and/or functional relations.

6. Representational aspects of microperspectives

Reference polarity. Every linguistic spatial relation is asymmetric in the following sense: One of its objects, the trajector or located object (LO), is *thematic* (a role which might derive from a question *Where is X?*), the other holds the role of a landmark or reference object (RO). With the two possible microperspectives of an implicit relation, there are two assignments of these roles to the source and goal object of a microperspective. I will represent this difference in the following by a feature *reference polarity (α refpol)* as specified in (6).

(6) a. +refpol: RO is source of a microperspective

b. –refpol: RO is goal of a microperspective

Note that this combination of perceptual and conceptual properties necessarily results in a subclassification of explicit spatial relations. Thus, we can predict that there may be two classes of linguistic spatial relations based on this subclassification.

Levels of representation. It is a characteristic feature of selective attention that it relates different levels of cognitive representation and processing. Accordingly, both perceptual and conceptual aspects are involved in the representation of microperspectives. On each level, the types of the attended objects and the type of the attentional change operation are specified, as well as the reference systems or frames with which the microperspective is associated. The latter may also involve the specification of a certain axis of a reference system, and the direction of the microperspective with respect to the axis. On the conceptual level, reference polarity will be marked.

Types of attended objects. There is abundant evidence that on the perceptual level, attention is attracted by discontinuities of the so-called "map of locations" (Treisman, 1988). Qualitatively, at least three types of attended objects (metaphorically, entities lying in the spotlight of attention) can be distinguished: whole visuo-spatial entities ("blobs"), boundaries between regions, or groups of blobs. These types will be relevant for further subclassifications of linguistic spatial relations.

Types of attentional change. There are three types of operations transforming one visuo-spatial attentional state into another: Attention can either be shifted, or it can be zoomed in or out. Conceptually, only shifts of attention between objects are relevant, regardless of the aspects of their visuo-spatial referents.

Reference systems. Levinson (1996) gives an overview over the different types of reference systems or frames in cognitive science. He presents a standardization of the different notions used in various disciplines and across cognitive modalities, and proposes to assume three different reference frames: the *relative* one, where the origin of the frame coincides with the viewer (usually the ego/speaker), the *intrinsic* one, where it coincides with the reference object (object centered and allocentric), and the *absolute* one where the reference frame is determined by aspects of the environment.

Despite the usefulness of a unified terminology, however, it makes sense to assign some aspects of reference systems to different levels of representation. For example, it has been shown (Kosslyn, 1994) that one has to distinguish between coordinate and categorial spatial relations, which corresponds to the assumption of spatial information being represented on the perceptual and conceptual level, respectively. Coordinate systems (retina-, head-, body- etc.) will therefore be placed on the perceptual level, while Levinsons concepts will be placed on the conceptual level (although they can also be viewed as metatheoretical descriptions of information represented on both levels).

Microperspectives´ relation to axes. Complementary to holistic reference systems, it has been proposed that the axes of so-called spatial frameworks (Bryant et al., 1992) are highly relevant for the characterization of spatial relations (and for the representation of spatial object knowledge, cf. Lang & Carstensen & Simmons, 1991). Accordingly, spatial axes (the vertical, the observer axis, and the left/right lateral axis) can be viewed as salient reference entities with respect to which microperspectives are categorized.

Congruence of axis direction and direction of microperspective. If microperspectives are related to spatial axes, then this relation can be further specified: the direction of a microperspective may either be congruent or incongruent with the direction of an axis. While the vertical and the observer axis are inherently directed,

directionality has to be imposed on the non-directed lateral axis. In any case, the congruence of both directions can be represented as a qualitative binary feature.

Taken together, these aspects of microperspectives can be represented in a frame-like non-linguistic structure shown in figure 3.

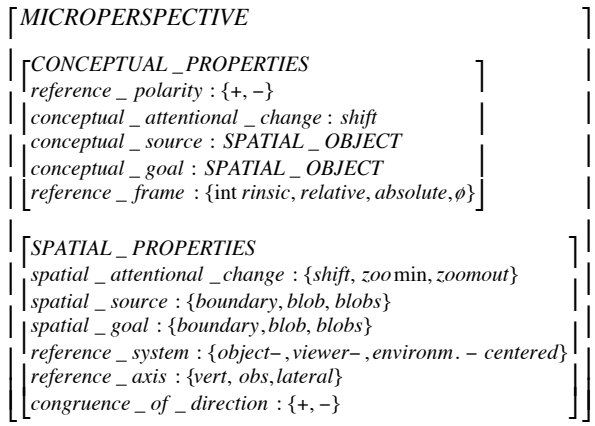


Figure 3: Schema for the representation of microperspectives

7. Localization as mental presentation

According to the classical view, localization is the central concept underlying the use of location expressions modelled by region inclusion. However, localization is at first a communicative phenomenon. It always depends on a question (called *quaestio*, cf. Klein & Stutterheim, to appear) that initiates or causes the description of the place of an LO. This quaestio is either stated explicitly by the hearer (*Where is LO?)* or is implicit in the speaker´s description of an object (e.g., *the LO behind the house*), thereby serving to comply with general communication principles („be informative“ etc.). In each case, it is the *speaker* who must provide an answer to this quaestio by localizing LO him-/herself. In contrast to most approaches which take an interpretation

perspective on localization, its investigation must therefore start with an analysis of the speaker´s cognitive processes during answer generation.

What does localization in terms of cognitive processes look like? I propose that it can be viewed as setting up a *mental presentation* of a spatial relation, where the construction of this presentation consists of a sequence of processing steps to be described in terms of basic cognitive operations. {2}

The following steps must be performed *necessarily* by the speaker (not considering the case where an answer to the quaestio is already available in propositional format, e.g. from information of the previous discourse) if an RO-centered spatial relation is expected to be the result of cognitive processing:

(A) Locating the LO. As a prerequisite for qualifying/describing the place of the LO, the speaker first has to locate the LO, either by finding it in the actual spatial context (by visual or haptic exploration) or by retrieving the relevant information from visuo-spatial memory. Locating the LO then means attending to and focussing a visuo-spatial referent that is instantiated in a processing medium (some part of working memory like the visuo-spatial sketchpad of Baddeley, 1982, the visual buffer of Kosslyn, 1994, the map of locations of Treisman, 1988, or the like). As to the imagery case, this corresponds to constructing and attending the image of LO instantiated in *a typical size* according to Kosslyn, 1978. Locating the LO thus results in setting up a mental presentation of that LO, having rather specific size, scale and resolution characteristics.

(B) Noticing a relevant RO. Having located the LO, an adequate and relevant RO must be found in order to establish a spatial relation. In cognitive terms, this corresponds to *noticing* an RO, that is attending to a visuo-spatial referent available in the medium which is categorized as RO. In the imagery case this is possible (at least in principle) because in constructing the LO´s image, other related information is activated via associative memory, and is (in part) instantiated in the medium. According to attention theory, these elements of the medium are processed in a way that the most salient element attracts attention (in that case,

away from the LO´s referent). In other words, selection of a salient RO *results* from some basic cognitive operations.

(C) Focussing the RO; Imposing reference frames. According to the standard approaches to localization, the RO is the center of a spatial relation. Here, this corresponds to focussing the RO´s referent, which has the same implications for RO as has step (A) for LO: establishing a „typical view" of RO and instantiating related information in the medium (via associative memory in the imagery case). Of course, higher order cognition must ensure that LO´s referent remains highly activated or gets re-activated correspondingly. Being the reference object, reference frames can then be imposed on RO´s referent. Most importantly, focussing may lead to changes in scale and resolution of the represented objective scene. Focussing the RO, too, thus results in setting up a mental presentation having rather specific size, scale and resolution characteristics.

(D) Directing attention back to the LO´s referent. Having set up the reference frame with RO´s referent at its center, attention must be directed back to LO´s referent in order to compute the spatial relation between RO and LO. Note that this results in the computation of spatial relations as predicted by standard approaches (RO centeredness, and possibly use of reference frames).

Both (B) and (D) are subject to what may be called „visibility constraint". As an axiom of attention theory, attention can only be directed to a visuo-spatial entity which is present (and salient) in the medium. From this it directly follows that the mental presentations in (B) and (D) must contain the referents of RO and LO, respectively, for attention to be directed to them. We can now these analyze these processing steps in the light of the previous discussions/context. There are three things to note:

First, in contrast to standard approaches to localization which include only centeredness of the RO, step (A) clearly shows that there is centeredness of the LO, too. This is important because several aspects of the mental presentation (size of objects, scale of represented scene) are heavily influenced by centering/focussing an

object. Second, a closer look at step (B) reveals that it corresponds to the microperspectivization of an implicit spatial relation between LO and RO thus establishing an explicit spatial relation between them. The resulting microperspective is negative referencepolar (–refpol) as the RO is its goal. Third, step (D) likewise corresponds to microperspectivization. In this case, however, the resulting microperspective is positively referencepolar (+refpol).

Thus, if the speaker´s cognitive processes are taken into acccount according to a localization-as-mental-presentation view, not only may the LO at some time have the role of a centered entity but also are there two different explicit spatial relations constructed in the course of localization. How can these unexpected observations be explained? Recall that we above presupposed that localization leads to a RO-centered relation (a +refpol relation, in our terms). Let us give up this constraint. We then have to consider the possibility that in the course of language production, a –refpol relation is selected and linguistically encoded. Could such a relation be expressed by a linguistic spatial term and if so, why should a –refpol relation be preferred to the more detailed +refpol relation?

The answer to the first question must be positive: Although being quite unspecific (representing little more than „if you perceive/conceptualize LO, then you perceive/conceptualize also RO“), a –refpol relation is fully specified as a spatial relation, given that the latter is conceived as representing an attentional operation between elements in a visuo-spatial medium. As to the second question, there may be restrictions on processing the steps (C) and (D). Among these possible restrictions are the following:

Time pressure. For various reasons (e.g. a complex spatial task), the cognitive system may be under time pressure while localizing LO. As steps (A) and (B) lead to an acceptable result, a solution would be to simply omit steps (C) and (D).

Dearth of more specific information available. Imposing a reference frame is not an effortless task. In the intrinsic case, the RO has to be inspected wrt. to ist parts giving clues about the typical

orientation/perspectivization. In the relative case, the position of the speaker must be determined and the spatial relation between speaker and RO has to be established. In the absolute case, the relevant features of the environment (global directions or the like) must be made available. Each of these subtasks requires information, either gathered from perception or retrieved from memory, which may eventually be incomplete or even lacking. In this case, the only option for the cognitive system would be to „fall back“ to the result of step (B).

Chosen level of granularity/specificity. Granularity phenomena are ubiquitous in cognition and language. One aspect of this is that granularity serves a speaker to adapt to the information needs of the hearer. This need not always be an option for a detailed level of description. For example, if a truck driver wants to tell his wife that after unloading he is now heading back home, he might well utter the sentence *I´m on the road again* (using a spatial description from which the relevant information can be inferred). In this situation, a more detailed description like *I´m 200m behind milestone 1137 on route 86, about 280 miles from home* would be quite inadequate. Thus, an unspecific spatial relation is sometimes preferred to a specific one (rendering the processing steps involved in the computation of the latter superfluous), depending on the chosen level of granularity.

Violation of the „visibility constraint“. As a necessary condition for step (D), the referent of the LO has to be present or „visible“ in the visuo-spatial medium for attention to be directable to it („visibility constraint“). This requirement may not always be met, however. In situations where the the RO is much bigger than the LO, focussing the RO may change the scale and granularity of the mental presentation in such a way that the LO is not „visible“ anymore. Again, the cognitive system may then „fall back“ to the result of step (B).

Summarizing the discussion about the linguistic expression of –refpol relations, we have found that they are not only suffiently specified but that it is also not always necessary or possible to build a corresponding +refpol relation. On the basis of a generalized “localization as mental presentation”-view we can therefore legitimately predict by non-linguistic argumentation that the two general classes of spatial relations (+refpol, –refpol) could be reflected in language.

8. Attention-based semantics

Preposition subclassification. A cognitive, attention-based semantics of linguistic spatial relations can now be specified by reference to the general schema of microperspectives shown in figure 3. It is proposed that the two classes of spatial relations distinguished by different reference polarity values in fact *are* reflected in language and that they roughly correspond to the well-known projective/topological distinction. According to this view, projective prepositions express positive reference polar (+refpol) relations and topological prepositions express negative reference polar (–refpol) relations. This fits with the role of the reference object as the "center" of a projective relation and explains the source orientation of these expressions (e.g., *away **from***) as well as the goal orientation of topological prepositions (*close **to***). Most importantly, however, it explains why the latter are different (as uncovered by the compatibility phenomena) without having to recur to metalinguistic principles (like the modification condition of Zwarts & Winter), functional relations or the like.

The mapping of spatial prepositions to spatial representations is done by specifying semantic entries that refer to instances of the schema, in which the value for the reference polarity feature is set and in which the semantic arguments are assigned their roles accordingly. The generic entries for projective and topological prepositions are shown in (7) and (8), respectively. Note that the referential variable for the microperspective (mp) is listed in the argument structure (which will be important below). As determined by the argument structure, x is the LO and y is the RO.

(7) **Semantic core of projective prepositions**

λy λx λmp [mp inst MICROPERSPECTIVE
& reference_polarity(mp)=+
& conceptual_source(mp)=y
& conceptual_goal(mp)=x

& spatial_attentional_change(mp)=shift {default}]

(8) **Semantic core of topological prepositions**

λy λx λmp [mp inst MICROPERSPECTIVE

& reference_polarity(mp)=–

& conceptual_source(mp)=x

& conceptual_goal(mp)=y

& spatial_attentional_change(mp)=shift {default}]

While the attention-based semantic entries for projective prepositions will not be presented here (cf. Carstensen, 2001), (9) shows the semantic specifications for topological prepositions *an* and *bei*. It is proposed that they differ in the ception of the reference object: For *an* to be applicable, a boundary of the RO must be attended (corresponding roughly to the 'close view' of Herskovits, 1986), for *bei* the whole RO must be attended (the 'remote view' of Herskovits). It can be easily seen that this leads to the characteristic extensional regions shown in figure 1.

(9) **Semantic entries of topological prepositions**

an (*at*, *on*, *by*): λy λx λmp [… & spatial_goal(mp)=boundary]

bei (*near*, *by*, *at*): λy λx λmp [… & spatial_goal(mp)=blob]

(10) presents the attention-based analysis of *in*, which is, contrary to common assumptions, not co-classified with *an* and *bei* (cp. *deep*/**near in*). Besides that, (10) also accounts for the examples like *knot in the rope* mentioned above as problematic for the classical view. {3}

(10) *in*: λy λx λmp [mp inst MICROPERSPECTIVE

& reference_polarity(mp)=+

& conceptual_source(mp)=y

& conceptual_goal(mp)=x

& spatial_attentional_change(mp)=zoomin]

Compatibility of distance and location expressions. The insensitivity of semantic compositionality to type differences of the involved constituents was the source of the modification problem in the classical view. In attention-based semantics, these differences are explicitly represented on the basis of the reference polarity feature, and they are made available for semantic processes via mp. Compatibility phenomena are explained by (in)congruence of the microperspectives provided by prepositions and those required by distance phrases (as modifiers), respectively. With the referential variable mp placed in the argument structure of prepositions, modification of location expressions by distance expressions can be formally represented as in (11). In contrast to (3) the distance predicate therefore applies to mp, and not to x. {4}

(11) a. λx λmp [LOCATION(x,mp)] modified by λy [DISTANCE(y)]

b. λx [LOCATION(x,mp) & DISTANCE(mp)]

9. Conclusion

In this paper it was shown that various phenomena of location and distance expressions call for an alternative view of localization in language whose characterization could start with the following words: ‘One purpose of locative descriptions is to express attention-based explicit spatial relations mentally present in the speaker´s mind (for the hearer to construct a corresponding mental presentation)’. While the classical view focusses on the search aspect of localization from the viewpoint of the hearer, and on RO-centered region inclusion, the more general “localization as mental presentation”-view emphasizes the role of explicit spatial relations based on

microperspectivizations of implicit spatial relations, for the conceptual representations of both hearer and speaker.

According to this attention-based approach, selective spatial attention must be regarded as a "missing link" in current theories addressing the relation of language and space. By selecting spatial entities for conceptual processing, attention acts as a device that perspectivizes implicitly represented spatial relations and at the same time modulates resolution and granularity in spatial representations. These aspects are ultimately reflected in language so that (crosslinguistic) variation in the inventory of spatial terms (cf. Bowerman 1996) can be explained by the (language-specific) differences in how these aspects are expressed by single linguistic items, and thus, in how object relations are perspectivized in a given language. An example is the linguistic distinction of the topological prepositions *an* and *bei* which can be explained by the option of whether the RO is attended to as a *whole object* in the mental presentation, or whether it is a *boundary* of the RO on which attention is focussed.

A new attention-based semantics was presented that solves some of the problems of traditional approaches to the semantics of spatial prepositions and their combination with distance expressions.

References

Aurnague, Michel, & Vieu, Laure (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*, Berlin: Mouton de Gruyter.

Baddeley, Alan (1992): Working Memory. *Science 255*. 556-559.

Behrmann, Marlene, & Tipper, Steven P. (1994): Object-Based Attentional Mechanisms: Evidence from Patients with Unilateral Neglect. In: Carlo Umiltà, Morris Moscovitch (eds.), *Attention and Performance XV : Conscious and Nonconscious Information Processing* (Cambridge, MA:MIT Press).

Bloom, Paul, & Peterson, Mary A., & Nadel, Lynn & Garrett, Merrill F. (Eds.) (1996): *Language and Space*. Cambridge, Mass.: MIT Press.

Bowerman, Melissa (1996): Learning How to Structure Space for Language: A Crosslinguistic Perspective. In: Paul Bloom et al. (1996), 385-436.

Bryant, David & Tversky, Barbara, & Franklin, Nancy (1992): Internal and External Spatial Frameworks for Representing Described Scenes. *Journal of Memory and Language 31*, 74–98.

Carstensen, Kai-Uwe (1992): Approaching the Semantics of Distance Expressions: Problems and Perspectives. In: Peter Bosch, Peter Gerstl (eds.*), Discourse and Lexical Meaning*, Workshop-Proceedings des SFB 340 "Sprachtheoretische Grundlagen für die Computerlinguistik", Bericht 30, 49-59.

Carstensen, Kai-Uwe (1995): Semantic and Conceptual Aspects of Local Expressions: Critical Remarks on the 'State of the Art'. In: Manfred Bierwisch, Peter Bosch (Hrsg.), *Semantic and Conceptual Knowledge*, Workshop-Proceedings of the SFB 340 „Sprachtheoretische Grundlagen für die Computerlinguistik“, Bericht Nr. 71, 117-126.

Carstensen, Kai-Uwe (1998): A Cognitive Science Approach to the Semantics of Spatial Gradation. In: Petra Ludewig, Bart Geurts (Hrsg.), *Lexikalische Semantik aus kognitiver Sicht – Perspektiven im Spannungsfeld linguistischer und psychologischer Modellierungen*, Tübinger Beiträge zur Linguistik (TBL 439), Gunter Narr Verlag (Tübingen), 73-94.

Carstensen, Kai-Uwe (2001): *Sprache, Raum und Aufmerksamkeit [Language, Space, and Attention]*. Linguistische Arbeiten 432. Tübingen: Niemeyer Verlag.

Carstensen, Kai-Uwe (2003): „Asymmetries in Spatial Semantics“. In: Claudia Maienborn (ed.), (A)-Symmetrien – (A)-Symmetries. Beiträge zu Ehren von Ewald Lang. Papers in Honor of Ewald Lang. Stauffenburg Verlag.

Coventry, Kenny R., & Carmichael, Richard & Garrod, Simon C. (1994): Spatial Prepositions, Object-Specific Function, and Task Requirements. *Journal of Semantics* 11, 289–309.

Garrod, Simon, & Ferrier, Gillian, & Campbell, Siobhan (1999): *In* and *on*: investigating the functional geometry of spatial prepositions. *Cognition* 72, 167-189.

Glasgow, Janice, & Papadias, Dimitri (1992): Computational Imagery. *Cognitive Science 16*, 355–394.

Hernández, Daniel (1994): *Qualitative Representation of Spatial Knowledge*. Lecture Notes in Artificial Intelligence. Berlin, Heidelberg, New York, Tokyo: Springer.

Herskovits, Annette (1986): *Language and Spatial Cognition. An interdisciplinary study of the prepositions in English*. Cambridge Univ. Press.

Herweg, Michael (1991): *Some notes on an and bei*. Working paper in the Doctoral programme in Cognitive Science. University of Hamburg.

Jackendoff, Ray (1983): *Semantics and Cognition*. Cambridge, Mass.: MIT Press.

Klein, Wolfgang, & von Stutterheim, Christiane (to appear): Quaestio and L-perspectivation. In W. Kallmeyer (Hg.) *Perspectivity and Perspectivation in Discourse*. Amsterdam 2001.

Kosslyn, Stephen M. (1978): Measuring the Visual Angle of the Minds Eye. *Cognitive Psychology 10*, 356–389.

Kosslyn, Stephen M. (1994): *Image and Brain*. Cambridge, Mass.: MIT Press.

Landau, Barbara, & Jackendoff, Ray (1993): "What" and "Where" in spatial language and spatial cognition. *Behavioral and Brain Sciences 16*, 217–265.

Lang, Ewald, & Carstensen, Kai U., & Simmons, Geoffrey (1991): *Modelling Spatial Knowledge on a Linguistic Basis*. Lecture Notes in Artificial Intelligence 481. Berlin, Heidelberg, New York, Tokyo: Springer.

Latecki, Longin, & Pribbenow, Simone (1992). On Hybrid Reasoning for Processing Spatial Expressions. *Proceedings of the European Conference on Artificial Intelligence* (ECAI 92), Vienna, Austria.

Levelt, Willem J. M. (1989): *Speaking: From Intention to Articulation*. Cambridge, Mass.: MIT Press.

Levinson, Stephen C. (1996): Frames of Reference and Molyneux´s Question: Crosslinguistic Evidence. In: Paul Bloom et al. (1996), 109-169.

Li, Jie (1994): *Räumliche Relationen und Objektwissen am Beispiel* an *und* bei. – Tübingen: Gunter Narr.

Logan, Gordon D. (1995): Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology 28*, 103–174.

Logan, Gordon D., & Sadler, Daniel D. (1996): A Computational Analysis of the Apprehension of Spatial Relations. In: Paul Bloom et al. (1996), 493-529.

Miller, George A., & Johnson-Laird, Philip N. (1976): *Language and Perception*. Cambridge, Mass.: Belknap Press.

Olson, David R., & Bialystok, Ellen (1983): *Spatial Cognition*. London: Lawrence Erlbaum Associates.

Peterson, Mary A., & Nadel, Lynn, & Bloom, Paul, & Garrett, Merrill F. (1996): Space and language. In: Paul Bloom et al. (1996), 553-577.

Pribbenow, Simone (1993): *Räumliche Konzepte in Wissens- und Sprachverarbeitung*. Wiesbaden: Deutscher Universitätsverlag.

Regier, Terry (1995): A Model of the Human Capacity for Categorizing Spatial Relations. *Cognitive Linguistics 6*, 63–88.

Talmy, Leonard (1983): How language structures space. In: H. Pick, Jr., L. P. Acredolo (eds.), *Spatial orientation: Theory, research, and application*. New York: Plenum Press.

Talmy, Leonard (1995): The windowing of attention in language. In: M. Shibatani, S. Thompson (eds.), *Grammatical constructions: their form and meaning*, Oxford: Oxford Univ. Press.

Talmy, Leonard (2000): *Toward a Cognitive Semantics*. Vol. I. Cambridge et al.: The MIT Press.

Theeuwes, Jan (1993): Visual Selective Attention: A Theoretical Analysis. *Acta Psychologica 83*, 93–154.

Treisman, Anne (1988): Features and Objects: The Fourteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology 40*, 201–237.

Tversky, Barbara (1996): Spatial Perspective in Descriptions. In: Paul Bloom et al. (1996), 463-491.

Winter, Yoad (2001): Measure Phrase Modification in Vector Space Semantics. In: K. Megerdoomian and L.A. Bar-el (eds.), *WCCFL 20 Proceedings* (Somerville, MA: Cascadilla Press), 607-620.
Zwarts, Joost (1997): Vectors as Relative Positions: A Compositional Semantics of Modified PPs. *Journal of Semantics* 14, 57-86.
Zwarts, Joost & Winter, Yoad (2000): Vector Space Semantics: a model-theoretic analysis of locative prepositions. *Journal of Logic, Language and Information*, 9:169–211.

Notes

---

1 I use the prefix „micro“ in order to keep this notion distinct from aspects of (representations of) observing a spatial scene which is usually associated with spatial uses of „perspective“ (cf., e.g., Tversky 1996). Besides that, microperspectivization corresponds in part to the processes of perspectivizing space in *micro*planning utterances (cf. Levelt 1989).

2 I deliberately use the term “mental presentation” in order to avoid the implications of using terms like “mental image” or “mental model” and to focus on the mere presence of a spatial relation in working memory. The spatial relation is “presented” to processes to operate upon it (categorization, instantiation) in the non-homunculus sense of being available to them. The term therefore complements Talmy´s (Talmy 2000) “ception” as a generalization of conception and perception.

3 As to functional relations, the position taken here is that they are not necessary parts of spatial semantics but that they nevertheless influence the *use* of spatial expressions. Functional relations constrain the possible range of implicit relations between two objects and may impose quasi-dynamic structures (e.g., the “representational momentum” of gravity or liquids and their inhibition represented by SUPPORT and CONTAINMENT, respectively) on them due to their representing physical principles and events. Because of the latter, re-(micro)perspectivizations are induced which prevent the plain-state-of-affairs (micro)perspectivizations and, accordingly, the use of the corresponding spatial expressions. This is the case with the well-known pear-hanging-in(to)-the-bowl example (e.g., Coventry et al. 1994) and the potato-under-the-bowl example (e.g.,

Herskovits 1986). In this sense, location control (cf. Garrod et al. 1999) accompanies or even determines explicit spatial relations but is not expressed by spatial prepositions.

4 A detailed analysis of gradation phenoma (which subsume distance phenomena in language) is presented in Carstensen (1998), and a discussion of several asymmetries in spatial language (including polarity phenomena) can be found in Carstensen (2003).

Abstract

Combining location and distance expressions for the description of spatial relations seems to be a natural thing to do in language. Recently, however, it has been observed that this combination is constrained and that standard approaches in the area of language and space fail to cover the critical phenomena (leading to the so-called *modification problem*). This paper presents an attention-based approach which offers a solution to this problem. It is based on the assumption that spatial expressions do not directly express aspects of spatial representations but that they express (micro)perspectivizations of spatial representations induced by spatial selective attention. The role of the latter as a "missing link" in the relation of language and space is exposed as being relevant for both linguistic differentiations and non-linguistic aspects (among them resolution and granularity in representations of space).