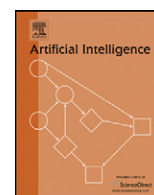


Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint


A linguistic ontology of space for natural language processing[☆]

John A. Bateman^{*}, Joana Hois, Robert Ross, Thora Tenbrink

SFB/TR 8 Spatial Cognition, University of Bremen, Germany

ARTICLE INFO

Article history:

Received 19 October 2008
 Received in revised form 24 May 2010
 Accepted 25 May 2010
 Available online 1 June 2010

Keywords:

Linguistic ontology
 Spatial language
 Natural language semantics
 Space
 Spatial knowledge

ABSTRACT

We present a detailed semantics for linguistic spatial expressions supportive of computational processing that draws substantially on the principles and tools of ontological engineering and formal ontology. We cover language concerned with space, actions in space and spatial relationships and develop an ontological organization that relates such expressions to general classes of fixed semantic import. The result is given as an extension of a linguistic ontology, the Generalized Upper Model, an organization which has been used for over a decade in natural language processing applications. We describe the general nature and features of this ontology and show how we have extended it for working particularly with space. Treating the semantics of natural language expressions concerning space in this way offers a substantial simplification of the general problem of relating natural spatial language to its contextualized interpretation. Example specifications based on natural language examples are presented, as well as an evaluation of the ontology's coverage, consistency, predictive power, and applicability.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The relation between language and space has long been an area of active research (cf. [152,74,106,27,28,154,111]). Human languages impose particular linguistic constructions of space, of spatially-anchored events, and of spatial configurations that relate in complex ways to the spatial situations in which they are used. Establishing tighter formal specifications of this relationship has proved a considerable challenge and has so far eluded general solutions. One reason for this is that the precise nature of the contribution made by spatial language has been conceived too simply. In much earlier and ongoing work, language is assumed to offer a relatively simple inventory of terms for which spatial interpretations can be directly stated. Examples of this can be found not only in accounts that focus on formalizations of particular tasks, such as path descriptions [112], scene descriptions [33], navigation and way-finding [158,163,140], but also in foundational work on the formal ontology of space [30,120,22], on qualitative spatial calculi [43,42], and on cognitive approaches involving, for example, image schemas [135,97]. In all of these approaches the principal burden of explanation is located within the non-linguistic formalizations pursued. This produces characterizations of spatial semantics that mirror the tasks and formal criteria addressed rather than the properties required for treating spatial language. As we shall argue below, such characterizations turn out to be ill-suited for dealing with the extreme flexibility of spatial language use observable in real contexts.

[☆] This paper has been improved substantially by the detailed critical comments of three anonymous reviewers and Thomas Stolz (Bremen). We also gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG) through the Collaborative Research Center on Spatial Cognition (SFB/TR 8) for the work reported.

^{*} Corresponding author.

E-mail address: bateman@uni-bremen.de (J.A. Bateman).

To bring the flexibility of spatial language use under control, we pursue a detailed formalization of what language itself brings to the interpretation of space. In particular, we consider language as contributing a structure of the spatial world that can be formulated as an ontology. This organization provides an additional layer of ontological information that formalizes the ‘semantic commitments’ entered into by any linguistic spatial construction. Such semantic commitments are intended to capture precisely the degree of formalization required to explain the linguistic options taken up while at the same time avoiding *over-commitment* with respect to the physical or conceptual spatial situations in which those linguistic options may be exercised. This then serves as an intermediate ‘interface ontology’ supporting mediation between linguistic forms on the one hand and contextualized interpretations on the other. We organize this intermediate layer around the central notion of ‘spatial modalities’, abstract categories that group collections of similarly functioning linguistic spatial constructions together without committing to particular contextual interpretations. For example, linguistic constructions involving the prepositions ‘at’, ‘near’ and so on are related to the spatial modality Proximal, which stands in contrast to Distal and is a superconcept of spatial modalities involved with lateral (left–right) and frontal (front–back) ‘projection’. The characterization of the linguistic expression “Bill is waiting at the post office” as expressing the spatial modality Proximal then makes commitments to a functional connection between the person waiting and the post office without (wrongly) committing to a spatial contextualization involving containment; both inside and outside the post office are still compatible with descriptions with ‘at’. An expression ‘near the post office’ brings in addition a contribution from the spatial modality Disjointness and so overlaps with, but is distinct from, ‘at’. The interface ontology as a whole provides a rich web of semantic inter-relationships; each of the spatial modalities we define brings with it semantic commitments that help constrain appropriate contextualizations. This paper motivates the design of the ontology in depth and provides extensive details of its use.

Constructing a view of spatial semantics as an additional layer of ontology in this way brings several advantages crucial for adequately capturing the relationship between language use and spatial interpretation. First, it supports the application of the full range of methods developed within ontological engineering and applied ontology in order to organize the information necessary in ways that conform to a strict and formally specified modeling style [65,67,148]. Second, it provides a suitable level of abstraction for dealing effectively with spatial language and for describing what linguistic expressions themselves bring to the interpretation process—something that has not been found possible when focusing on linguistic elements as isolated terms. And third, it allows the relationship between linguistic expressions and spatial interpretation to be recast as a particular case of ontological alignment, or mediation, whereby two or more distinct ontologies are brought into a formal relationship [89,100,99,80,98]. Combining these considerations establishes a formally robust and well grounded framework from which to consider the full flexibility required for dealing with the mapping between language and space.

A distinct, well-defined and empirically-motivated layer of semantics for spatial language of this kind has direct applications for many currently relevant tasks involving spatial language. These include attempts to support communication via natural language with the human users of Geographic Information Systems [21,118,97,24], of context-based services [20], and of devices operating *in* space, such as situated robots [137,122,96]. In addition, the growing demand for automatic enrichment of textual data with spatial annotations in the context of the semantic web raises precisely analogous issues [116]. In all of these areas, solutions to the language–space mapping problem are urgently required. We also suggest that the organization we propose is equally relevant as a contribution to the *linguistic* discussion of space, particularly concerning how the interpretation of spatial language can be managed.

We organize our discussion as follows. In Section 2 we establish precisely why we consider spatial language to demand its own layer of ontological description. We then present in Section 3 two examples of linguistic spatial descriptions illustrating the general spatial categories of our linguistic ontology and their connection both to natural language constructions and to formal contextualized interpretations. In Section 4 we turn to the detailed structure and definitions of our linguistic ontology, including illustrative examples. Section 5 discusses issues concerning applicability and use of the ontology, and in Section 6 we address evaluation—itsself now an area of increasing importance for ontological engineering in general. In Section 7, we conclude with a summary of what has been achieved and consider future research tasks.

2. The case for a linguistic ontology for space

Our claim is that the processing of authentic spatial language requires a particular kind of knowledge to be captured—knowledge concerning the range of spatially-related meanings that language itself, or languages themselves, construct. This knowledge and its organization have been dealt with inadequately hitherto despite the fact that much is now known concerning the linguistic construction of space. Linguistic investigations of spatial semantics, particularly approaches from Talmy, Langacker, Vandeloise, Bierwisch, Lang, Levinson and others [152,155,106,25,105,161,110], have revealed broadly similar mechanisms at work across cultures and types of language use and a considerable body of work has emerged building on these foundations. Detailed studies of spatial language use can be found in contexts ranging over formal semantic interpretations, psychological studies, dialogue analysis, computational modeling and many more [128,159,160,36,157,744]. The insights gained in this tradition provide the starting point for the formalization we propose. It is of particular importance to be clear about just how substantial this previous work is and, for this reason, our bibliographical references are extensive—although still by no means exhaustive. So far, few attempts have been made to bring this substantial body of information together within a unified account supportive of automatic processing. We propose here for the first time a comprehensive approach to combining the various contributions that also relates them to non-linguistic spatial interpretations.

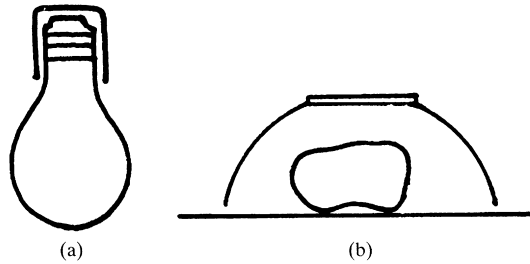


Fig. 1. Examples of problematic spatial containment configurations from Herskovits [74].

The fundamental question at issue has been framed succinctly by Bierwisch: ‘How much space gets into language’ [26]? We propose for this a *linguistically-motivated* ontology, or ‘linguistic ontology’ for short, that provides a specification of an encapsulated layer of ontological information motivated solely by the requirements of linguistically-expressed spatial meanings. This provides a new layer of organization for capturing the contribution of language to spatial interpretation that is free of non-linguistic, contextually-dependent additions. We thereby decompose and modularize the problem of interpreting and producing spatial language by ‘stratifying’ three ways: (i) lexicogrammatically, (ii) according to a shallow semantics, and (iii) by contextualized specifications—all three of which are formally distinct. The application of ontological engineering methods then offers considerable benefits for teasing apart the respective contributions made by these essentially distinct knowledge sources.

We motivate the use of a linguistic ontology intermediate between linguistic expressions and their contextualized interpretation by briefly considering three perspectives relevant to the modeling of spatial language: we need to address (a) the linguistic phenomena of spatial language use, (b) the formalization of spatial language interpretations, and (c) the computational instantiation of processing schemes for natural language involving space. In all three domains there is striking converging evidence in favor of characterizing the relative contributions of spatial language, particularly spatial semantics, and domain or task descriptions of space in the manner we suggest.

2.1. Evidence from linguistic usage

Early work in the semantics of spatial language began with a predominantly *geometric* perspective that still exerts considerable influence today. Under this view, the semantics of a spatial expression is related quite directly to a spatially-specified situation. For example, the semantics of a prepositional phrase involving the English preposition *in* might be given using a two-place predicate of the form:

$$in(x, y)$$

interpreted to mean that some spatially-locatable entity x in the world is contained geometrically within the boundaries of some spatially-identifiable region y . Problematic with this account, and all modifications of it, are the basic facts of linguistic usage. Extensive illustrations of this have been given by Herskovits [74]—one of which concerned with the preposition ‘in’ is presented in Fig. 1. In the situations depicted we cannot say, on the right-hand side, that the potato is ‘in’ the bowl even though it is, in some sense, completely contained geometrically, although, on the left-hand side, we *can* say that the bulb is in the socket even though it is patently not contained.

The kind of spatial meaning involved with ‘in’ is termed a *topological* spatial relation. Such relations generally locate an entity with respect to some specified region but do so in a way that makes it difficult to identify the regions that are applicable in any particular case. Although a ‘naive’, or pre-theoretical, consideration might suggest relations based on geometry (e.g., containment for ‘in’), this leaves many common instances of use to be considered as exceptions. Commonly proposed meanings for ‘in’ include at least [74,55,39,24]:

- geometrical containment,
- containment within a concavity (‘in a cup’),
- containment within a containing surface (‘a crack in the glass’),
- interposition among elements (‘in a forest’),
- location within/among elements of an aggregate (‘in a town’, ‘within/among buildings’).

These all involve some notion of ‘containment’ but precisely *what* is contained *where* is quite varied. Considerable effort has been expended in trying to bridge the gap evident between geometric descriptions of spatial situations and the kinds of meanings listed but, in general, the problem remains unsolved.

Bennett and Agarwal term these kinds of alternatives *modes of locating and hosting* and simply accept them as alternative meanings requiring alternative formalizations [24]. This ‘gets the job done’ as far as moving forward on spatial formalization is concerned but does not go far towards explaining linguistic usage. Indeed, the situation is even worse in that each of

these modes of locating and hosting can be driven further in ways that are then rendered more or less ‘metaphorical’. In a sentence such as (1), for example,

(1) I am in the tree.

Herskovits suggests that we can build on geometrical containment by adopting the three-dimensional convex hull of the tree rather than the tree itself. Whereas in a sentence such as (2),

(2) I am in Bruxelles.

Bennett and Agarwal propose that in this case the region for a containment relation is ‘footprint’ containment, whereby the two-dimensional ‘footprint’ of the spatial extent of the speaker (‘I’) is taken to be a spatial part of the two-dimensional ‘footprint’ of the spatial extent of Bruxelles. This is clearly an adequate analysis neither of example (1), nor of the examples depicted in Fig. 1.

In short, we arrive at a proliferation of formalizations for which it may be possible to *force* an interpretation in terms of some topological inclusion relationship but the main problem is actually deciding on which regions may be playing a role or not. If this is left sufficiently free, the geometric account is essentially empty—*some* region will always be constructible *post hoc*. The source of explanations is therefore located within contextualized interpretations without telling us precisely how such contextualization might be reached.

Another class of spatial meanings is described in terms of *projective* relations, such as ‘in front of’, ‘left of’, etc. These are more complex than topological relations in several ways. Their interpretation relies on the additional step of fixing an underlying *reference system* [110], defined relative to the speaker or hearer, or with respect to some object’s intrinsic or imposed orientation properties (‘my left’ vs. ‘your left’, ‘the front of the church’, or ‘in front of the rolling ball’, etc.). Unless the reference system is known, projective spatial expressions cannot be related unambiguously to spatial context. Nevertheless, it is precisely this crucial information that our English and German data show to be most commonly left implicit in natural dialogue, requiring resolution during interaction or from context [19]. Providing maximal support and constraint for this resolution is then extremely important for effective spatial language processing.

The problems that arise when attempting to fix the spatial import of projective relations are very similar to those of topological relations. We can show this drawing on the most commonly used formalizations proposed for projective relations, those of *spatial templates* [37,113], or *field potentials* [64,129]. These formalizations build on the observation that projective terms appear to identify general directions rather than precise orientations. For example, ‘to the left of X’ picks out a general direction towards the left but is not restricted to a 90° angle. Use of projective terms therefore suggests a *graded applicability structure*, whereby one direction may be a ‘best’ exemplar (a ‘focal axis’) but other directions are also acceptable to the extent that they approximate this best case. This can be captured by defining for each projective relation a probability field that reflects the likelihood of application of a term given a particular spatial position. This approach is applied in computational approaches to robotics and perception where a robot may need to interpret linguistic descriptions (particularly expressions containing spatial prepositions) to follow particular paths or locate objects [136,90,150,124,91]. More complex expressions (e.g., ‘in front and to the left of X’) are generally interpreted by defining compositions of the individual fields involved.

The graded probability of location obtained within such accounts often matches applicability judgements made by speakers with respect to spatial situations relatively well. But there remain problems, most significant of which is the basic issue of just how the potential field relevant for a particular communicative situation is to be determined. Potential fields need to be parameterized in various ways (e.g., by being stretched or relaxed according to relative size and orientation [166,74], with respect to potential distractors [91], and according to the selected reference frame), but the precise motivations for that parameterization are often context specific or *ad hoc*. This is exactly analogous to the problem with topological spatial terms: just as it was unclear what regions are called for, here it is equally unclear just what kind of field is appropriate.

Natural language usage is, moreover, substantially more varied than a straightforward parameterized selection of applicability fields would suggest. Consider, for example, the interpretation of even seemingly simple linguistic expressions involving *left*. A common interpretation compatible with potential fields would be in terms of an area of high probability along an axis situated 90° counter-clockwise with respect to some reference object plus an orientation. However, in different contexts ‘left’ can equally well denote the entire ‘left-hand’ half-plane, a reorientation by an angle of contextually determined size or a redirection of movement, i.e., the expression *turn left* may denote a change of orientation (on the spot), a change of movement direction without a reorientation (as, for example, in a sideways motion or with non-oriented entities such as balls), or both (intuitively the normal case in route instructions). Worse, in the case of a street network, ‘turn left at the next junction’ has little to do with a focal axis: in fact, the field collapses according to the actual possibilities provided by the road network. If there are several roads on the ‘left-hand side’ it is not enough to pick the one that is closest to a 90° angle and to call that the ‘left turn’, since all the turns on the left are ‘left turns’ in some sense; a more discriminating description is required [93].

Field potentials are also used for other spatial relations with similar accompanying problems. For example, one model of the spatial expression ‘at X’ might be as a circular field centered (i.e., with highest probability) on the element X. Or, a model of the expression ‘along X’ might be a field that follows the shape of the spatial element X [167]. But in each case, just what

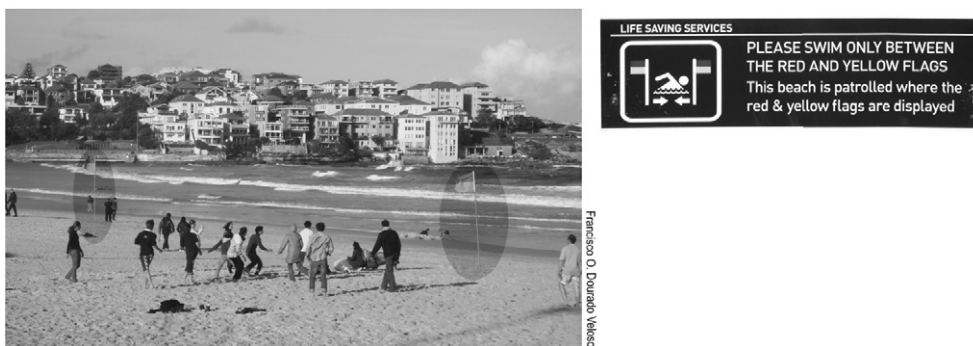


Fig. 2. Example of a seriously displaced spatial relationship taken from a sign at Bondi Beach, Sydney.

shape of field is to be picked as well as how it is to be positioned with respect to the current spatial situation represents a difficult interpretative decision in its own right. One last illustration of this is given in Fig. 2; this shows a repositioning of a potential field that is extreme, while still remaining a straightforwardly everyday example of spatial language that is unlikely to strike any but the most literal of interpreters (e.g., us and the computational systems we construct) as strange. The example concerns the spatial relationship ‘between’; if asked without any context what the appropriate potential field for describing the spatial situation corresponding to the spatial expression ‘between X and Y’ might be, it would be reasonable to expect that points nearer to a line connecting X and Y would receive a higher probability than points lying further away. The figure shows, however, that the information given in the explanatory sign (‘swim only between the red and yellow flags’) intends to pick out a completely different spatial situation and, what is more, ‘effortlessly’ succeeds in doing so—violating on the way a ‘universal’ hypothesized by Zwartz that the spatial interpretations of simple prepositional phrases necessarily contain their reference objects [168, p. 80].

Such examples show that relating a linguistic term to a probability field conflates at least two steps: the selection of a contextual interpretation and the generation of a field appropriate to that contextual interpretation. But, as with topological regions, it is precisely this first stage that is the main problem facing spatial language interpretation and production. Discussions of spatial terms often fail to separate these two stages and so already incorporate particular contextualized interpretations in their consideration of the linguistic semantics involved. With respect to projective terms, such discussion leads to the misconception that graded applicability structure is an integral part of some spatial term’s semantics [74], whereas it is actually a property of the contextualization [157]. It is only when we have the contextualization that we can go on to identify *which kind and shape* of probability field may be required and it is only there that we can observe graded applicability. But to reach this stage, we must first unravel the linguistic contribution to the contextualization process.

A third area of difficulty with natural language usage is the considerable ‘ontological flexibility’ that spatial terms exhibit. Depending on their context of use, particular spatial terms take on quite different properties. One classic statement of this was given by Hobbs over a decade ago. Consider the ontological status of a spatial entity such as 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” [78, p. 820].

These quite different ‘conceptualizations’ give rise to correspondingly different linguistic descriptions, or grammatical *syndromes* as we will characterize them below—such as ‘along the road’, ‘on the road’, ‘in the road’, and so on. Each of these possibilities commits to a certain range of properties for the object described, but those properties are only indirectly related to ontological properties of the entity in the world. Indeed, even the selection of a lexical item, such as ‘road’, cannot be taken as a neutral labelling of reality. There may be considerable uncertainty as to where the limits of applicability of a term might lie and to what follows from such an application once made; this problem has received considerable study in the area of geographic entities [23,147]. In general, the diverging *identity criteria* necessary for the entities picked out by linguistic classifications and by entities ‘in the world’ demand an ontological separation [31,68].¹

Language usage itself requires us, moreover, to distinguish between at least ‘objects’ and the ‘places’ where such objects may be located [103]. Several grammatical tests support this distinction. Asher, for example, suggests use of a grammatical alternation involving (in English) *in* or *at* versus *inside* [5]. This test relies on the fact that it is quite dispreferred to say that one is ‘inside’ a region; this motivates the contrasting acceptability judgements observable in examples (3)–(4):

- (3) The tractor is in the field.
- (4) ??? The tractor is inside the field.

¹ For a rather different approach to some aspects of this issue, see Pustejovsky’s use of ‘dotted types’ within the generative lexicon [133, p. 334].

A 'place' may be brought into discussion by referencing an object but it remains a question of interpretation to determine just what that place is.

It is also clear that language usage demands an account in which geometric modeling is often not the primary consideration. One extension beyond geometric information now finding wide acceptance is the inclusion of *functional* notions [161]. Functional relationships obtain whenever the *use* or *behavior* of one entity depends in some way on another. There is now strong evidence that a broad range of spatial terms are crucially sensitive to functional features [45,36]. For example, acceptability judgements for the use of 'above' when referring to a tube of toothpaste geometrically 'above' a toothbrush vary systematically according to whether the precise placement is supportive of putting the toothpaste on the brush or not; moreover, for some particular linguistic terms, it is the functional information that appears to be a stronger conditioning factor of use whereas, for others, it is the geometric contribution that appears stronger [46].

Common functional notions include 'control' and 'support'. These deal succinctly with the problematic uses of 'in' shown in Fig. 1 above: the bulb is 'in' the socket because it is 'functionally controlled' by this placement—if we move the socket the bulb moves, if we remove the socket, the bulb falls, etc.; similarly, the potato is not 'in' the bowl because it is *not* controlled by the bowl—if we lift the bowl the potato stays where it is. It does not appear plausible at this time to reduce the functional to the geometric, or *vice versa*. Indeed, language usage in general is most responsive to *perceived* (interpretation) or *claimed* (production) functional relationships—regardless of their geometric support.

In all three of the areas of linguistic usage discussed—topological terms, projective terms, and the ontological flexibility of spatial categorization—we find considerable subtlety being introduced by the need to consider language. There are distinct *kinds* of information, or entities, being drawn upon which may not be evident when considering a treatment of space independently of language. The complex and situation-dependent relationship of these entities to spatial interpretation calls for their own layer of ontological modeling.

2.2. Formalization evidence

In the previous subsection we have seen that language use is very flexible. In this subsection, we focus on some of the problems that this flexibility brings for formalization—at this stage still independently of whether a computational instantiation is being targeted. Formalizing the meanings entailed by this flexibility clarifies further the difficulties involved. Particularly problematic is the fact that specific linguistic terms have been taken to require very different logical formalizations in different circumstances. This follows directly from the sheer diversity of Bennett and Agarwal's 'modes of locating and hosting' mentioned above. As they write:

"Because of the ambiguity of natural language, there will not be a definite mapping between natural language terms and elements of this semantic theory. Rather, we shall find that each natural language term has a number of distinct senses, corresponding to different ways in which the interpreted entity can be interpreted within our semantic framework" [24, p. 80].

But from a linguistic perspective, allowing unrelated meanings within a lexical item is less than satisfactory. It is often a sign of missed generalizations and complicates processing. Here we wish to ensure that we do not lose the strong sense in which shared grammatical constructions share important aspects of their meaning.

The interposition of a knowledge organization mediating between linguistic form and contextualized interpretation, i.e., our linguistic ontology, attempts to capture these shared aspects directly so as to provide a more parsimonious framework for integrating the 'distinct senses' of spatial expressions proposed in the literature. These distinct senses are often only required because they are already attempting to model *contextualized interpretations*. In effect, each formalization has already committed to a particular usage in context without telling us how that might have been reached. Accounts are then forced to simply include further alternate logical forms within the lexicon entries involved. Building a more abstract, linguistically-sensitive layer of formalization into the account allows us to move beyond this in order to be more supportive of flexible spatial language usage that is nevertheless constrained to lie within the range of variation that naturally occurs.

In many respects, our formalization approach is analogous to the direction started by Bierwisch. Bierwisch [26, p. 36], for example, presents the following two examples illustrating considerable diversity in meaning for very similar constructions:

- (5) He left the institute an hour ago.
- (6) He left the institute a year ago.

In (5) we appear to be dealing with a change of location; in (6), however, the meaning appears to be a change of social affiliation. Bierwisch's method of dealing with this diversity is to interpose a semantic representation that is neutral between the distinct contextualized readings. In this case, a single semantic representation is postulated, centered around a generative semantics-style predicate-argument expression such as:

[x DO [BECOME [NEG [x AT y]]]]

The central 'spatial' component involving the semantic prime 'AT' is then kept neutral between *x* being spatially located proper within the physical space of the institute and *x* changing their social affiliation with the abstract entity constituting

the institute. This offers a straightforward example of factoring out a context-independent semantic representation in order to provide a suitable starting point for contextualization.

In several more recent approaches to the formalization of linguistic spatial expressions, we find this separation of concerns not being followed. It is therefore interesting and relevant to consider briefly whether the problems of usage of linguistic spatial terms introduced above are appropriately included in these approaches. Kracht [94], for example, presents an extensive, typologically-motivated proposal for the semantics of many spatial expressions in terms of a presumed universal semantically transparent phrase structure of the form:

$$[MP\ M\ [LP\ L\ [DP\ Landmark\]]]$$

This represents structurally an assumption that locative expressions consist of two ‘layers’, one of which Kracht terms the *configuration*, the other the *mode*. Configurations capture mutual static spatial relations between objects and are expressed as ‘location phrases’ (LP); modes capture movements with respect to those configurations via ‘mode phrases’ (MP). An important benefit of Kracht’s approach is that the semantics of such phrases, combinations of such phrases, and their embedding within larger syntactic contexts are all constructed in a strictly compositional and well-specified fashion. This is a desirable property for any account and is certainly one which we attempt to maintain in our own approach. Particularly telling, however, are the consequences for compositionality that Kracht’s basic assumptions give rise to.

Kracht takes the semantics of spatial expressions to be temporally parameterized spatial neighborhoods in three-dimensional space. In particular, the lexical semantic content of spatial terms, such as prepositions, is provided by a collection of specially defined localization functions whose task is to map given spatial neighborhoods to other *geometrically related* spatial neighborhoods. For example, the semantics of the preposition ‘in’ under this account is a function that requires an object x and a time t and which produces the set of regions that are contained in the inside of x at that time t . Now, while this is without doubt a possible and appropriate contextualized interpretation of *some* uses of spatial expressions, it is by no means clear that this is an appropriate semantics for the linguistic expressions themselves—for all of the reasons of flexibility sketched in the previous subsection. The consequences of this for compositionality are, however, substantial and Kracht himself notes that he explicitly wishes to exclude examples of the kind Herskovits discusses; in fact:

“...there are many instances where [the semantics of localizers we have given] fail to do justice to our intuitions. Moreover, there invariably are borderline cases. For example, when is some object x *at* an object y , and when is it *on* y ? ... It is difficult to imagine how a strictly compositional account would handle this problem, and we shall have to leave it aside” [94, p. 191].

Thus, compositionality is only maintained in this account by relying on the natural compositionality of geometric operations on regions: linguistic compositionality has been made dependent on a particular (contextualized) spatial interpretation, one which, as we have seen, is itself problematic.

A similar difficulty arises in the combinatory categorial grammar approach to spatial expressions proposed by Francez and Steedman [58]. This account, although couched within a very different formal and linguistic framework, makes the same assumption that an appropriate semantics for spatial expressions is a specification of the regions of three-dimensional space that those expressions pick out. Here, as is usual within combinatory categorial grammar accounts, the semantics of larger expressions is constructed via function application over the semantics of component expressions as determined by the internal structure of complex syntactic categories. Francez and Steedman add to this semantics explicit spatial and temporal ‘contextual variables’ which, during function application, collect additional constraints via so-called ‘shift-functions’ [58, p. 394]. In the case of spatial expressions, the role of these functions is to take some spatial region and to modify, or ‘shift’, this to produce another region related in some particular specified way to the originating one.

The precise semantics of these shift-functions is not Francez and Steedman’s main focus—they are primarily concerned with achieving an appropriately compositional account that adequately reflects the role of context. In this respect, there is much to be taken from the mechanisms they propose. However, it is clear that these shift-functions are, again, being seen primarily as involving geometric operations. This can perhaps be explained in relation to the origins of their approach in Pratt and Francez’s treatment of *temporal* semantics in terms of contextual variables [131]. In that account, the semantics of temporal expressions is related directly to temporal interval logics, enabling the interpretation of linguistic temporal expressions to proceed in step with the operations these logics support. Whereas this treatment is very appealing for treatments of temporal language, the use of *spatial* language shows considerably more flexibility and it is by no means clear which ‘logics’ (and there are many (cf. [42,1])) may appropriately play the role that is played by temporal interval logics in the case of time. It is clear, however, that Euclidean three-dimensional geometry does not provide a sufficient basis; this situation therefore needs to find appropriate consideration within any account that seeks to cover a broader range of natural spatial language usage.²

The direction we follow in our development of a linguistic ontology for spatial language accordingly draws on a further class of recent approaches to the semantics of spatial language. These are approaches which use linguistic evidence in order

² More discussion of Kracht’s and Francez and Steedman’s accounts with respect to our requirements for spatial contextualization is given in Bateman [13].

to motivate abstract specifications of spatial semantics that cover just those distinctions that language appears to require, rather than relating this immediately to contextualized interpretations in a three-dimensional space. This follows further the ‘two-level’ approach of Bierwisch introduced above, but begins spelling out in considerably more detail the linguistic side of the spatial equation.

An early detailed account of this kind was Eschenbach’s [53] formalization of spatial expressions such as ‘left’/‘right’, ‘in front of’/‘behind’, etc. In this framework, an *abstract* geometric structure covering the semantic commitments of language is imposed on the *concrete* spatial situation, depending on the binding of some specific, well-defined contextual anchors. The linguistic semantics is then couched in terms of abstract algebraic structures, with their own entities and properties—for example, that required for ‘left’/‘right’ involves half-planes, that for ‘in front of’/‘behind’ involves oriented half-lines—which may then be ‘anchored’ into particular contexts in a flexible but well-defined fashion; we have discussed this approach and its relation to linguistic ontology elsewhere [19]. Another well developed account in this spirit is the work by Zwarts on the semantics of location and path expressions [168,169]. Employing several distinct kinds of linguistic evidence (most interestingly, aspect), Zwarts [169] develops in a style analogous to that of Eschenbach an algebraic theory that yields appropriate properties for path descriptions constructed compositionally during the interpretation of prepositional phrases, combinations of prepositional phrases, and prepositional phrases combined with other spatially-relevant material in their respective verb phrases. Again, these constructions must be seen first and foremost as abstract spatial interpretations of linguistic spatial *descriptions* which may stand in more or less complex relationships with actual spatial situations.

We consider the moves being made in this direction as offering strong evidence in favor of the kind of modularity that we pursue here. Distinguishing a linguistic spatial semantics from contextualized spatial interpretations appears to be a beneficial, and perhaps even necessary, step towards getting the full flexibility of spatial language usage under theoretical and formal control. The main task facing the development of such a semantics further is then to characterize in detail what kinds of abstract spatial configurations can be motivated linguistically, what their interrelationships and properties are, and how they interact with other aspects of semantics and context.

2.3. Natural language processing concerns

The interpretation of spatial language raises similar general issues to those involved when trying to relate generic natural language processing components to the requirements of any specific domain. The basic premise we motivate here is that it is useful from several perspectives to specify a level of representation that is intermediate between natural language expressions on the one hand and formally specified characterizations of spatial situations on the other. A separation of information into distinct modules of this kind is reminiscent both of earlier ‘two-level semantics’ accounts that have been adopted in several natural language processing contexts [105,75] and of notions of ‘quasi-logical form’, originally advocated in systems such as the Core Language Engine [2]. Our own use of a two-level architecture goes back to the *Penman Upper Model*, originally developed within the Penman text generation system in the mid-1980s [117,16,12]. This was, to our knowledge, the first approach to two-level semantics explicitly formulated as an ontology. Particularly within natural language *generation*, it was recognized relatively early that organizing ‘domain knowledge’ in a way that reflected its expression in natural language would more readily support generic natural language generation applications. The Penman Upper Model was accordingly defined as an interface between application knowledge and linguistic knowledge, expressed as what we would nowadays term a ‘lightweight ontology’ using the knowledge representation system LOOM [115].

The introduction of a further layer of domain-independent semantics between syntactic analysis/generation and domain knowledge is a move now being suggested, apparently independently, as a component of several distinct approaches. Within a parsing and interpretation context, it has been argued that such a semantic layer improves portability and re-use of components within dialogue systems [51]; within a generation context, it has been argued similarly that compositional semantics needs characterizations that capture how language decomposes entities and that this is, again, independent of domain-specific organization [149]. Within the spatial domain, we also see language-motivated characterizations proposed by Mavridis and Roy as a kind of ‘parsing’ of “situations into ontological types and relations that reflect human language semantics” [122]. Here, just as in our case, the relationship to language is intended to support automatic natural language processing, while the relationship to situations and ontological types is intended to ease their formal interpretation and contextualization.

A further motivation for imposing this additional layering, or stratification, of ‘ontological’ representations when considering language processing is that it provides a more appropriate modularization of concerns. Without stratification there is a tendency to equate existing foundational ontologies, such as DOLCE [119], SUMO [125] and so on, with natural language semantics directly. Such treatments can be sophisticated [63,40], or more straightforward as in the direct linking of SUMO categories with WordNet synsets [126] or in the inclusion of semantically-motivated categories as a proper part of the ontology as a whole, as in the Sensus ontology of the DARPA knowledge-sharing initiative [151] and the more recent Omega effort [84,130], both of which build on developments of the Penman Upper Model mentioned above. Problematic in all these cases is that they do not address the ontological flexibility of natural language terms that we discussed above with respect to the quote from Hobbs and, as a consequence, mix components of very different ontological statuses. A representation of natural language semantics within a foundational ontology, for example as proposed for DOLCE by Cimiano and Reyle [40] or for SWINTO (a combination of DOLCE and SUMO) by Oberle et al. [127], then commits to an inappropriate rigidity of interpretation. Similar criticisms hold for links between lexical organizations, such as FrameNet, VerbNet, WordNet, etc. and

foundational ontologies of any kind. The flexibility of ‘ontological’ classification requires that entities in the world receive diverse classifications according to the requirements of the discourse context. When bringing foundational ontologies and linguistic information together, therefore, it is necessary to pay particular attention to the ontologically non-rigid nature of linguistic classification.

It is equally necessary to avoid importing into computational system design over-restrictive assumptions from theoretical accounts. For example, Jørgensen and Lønning develop a Minimal Recursion Semantic analysis component for computational use that is embedded within Head-driven Phrase Structure Grammar and which implements fairly directly the account of Kracht summarized above [88]. Based on Kracht’s assumptions, they assume that it should be reasonably straightforward to move from their underspecified semantics to contextualized interpretation—however, as argued above, this will only hold for geometrically transparent usages. It would be highly desirable to maintain the overall framework for interpretation that they develop, while at the same time weakening the implicit reliance on geometry.

Our approach to meeting these requirements draws, as mentioned, on the course of development begun with the Penman Upper Model. Here, the basic idea was to provide an ontology of categories and relations sufficient for driving all of the decisions required by the linguistic components present in the system. Knowledge from particular domains was then made accessible to the linguistic components by formally relating (at first in terms of subsumption or inheritance) the concepts from any specific domain to the concepts provided by the Upper Model ontology. Any domain-specific concept then inherited appropriate ‘methods’ for linguistic expression [12, p. 57] and semantic specifications could employ domain knowledge concepts freely, safe in the assumption that the linguistic components would know how to express those concepts via the methods associated with the linked Upper Model concepts. Crucial to the design of the Upper Model was a commitment to including distinctions only when they are motivated by specifiable contrasts in grammatical form. For this reason, we consider an Upper Model to be a linguistically-motivated ontology representing that portion of the semantics of a natural language that finds expression in that language’s grammar; similarities can usefully be drawn here with Jackendoff’s *criterion of grammatical effect* [87, p. 13]. This guarantees that the relation to linguistic form is known—thereby supporting its use in generation and analysis—while also moving the description towards semantics.

Basing design on grammatical evidence rather than lexical organization differentiates the Upper Model from other, superficially similar systems that have been developed on lexical grounds, such as FrameNet, VerbNet, WordNet and systems descended from these [85]. We consider the distinction between lexical semantics and grammatical semantics central because lexical items tend to be too idiosyncratic in their bundling of semantic properties to reveal generic semantic configurations.³ The Upper Model is therefore more closely related to approaches that adopt an explicit orientation to *grammar* and *grammatical distinctions* for linking with semantics as now suggested across a very broad range of frameworks, from generative grammar [109,87], through cognitive approaches [61], to social-functional accounts [70]. The degree to which it is possible or desirable to bring the lexically and grammatically motivated organizations together remains an open issue at this time.

Our main motivation for calling an organization of this kind an ontology then lies in the role we attribute to language of structuring, or schematizing, experience. In many respects, just as any ‘reality’ is necessarily filtered through our perceptual systems, it is also filtered through the language system—at the latest whenever we wish to communicate (‘thinking for speaking’ [144]). We can therefore characterize this linguistically constructed reality in much the same way, and applying similar principles, as adopted when characterizing other constructions of reality.

The concrete starting point for the work described in this paper is a particular extension to the original Upper Model design made in the mid-1990s in order to meet new application demands. This involved both broadening the range of languages addressed and incorporating some generic semantic principles developed for ‘linguistically motivated’ ontologies [70]. The result was the *Generalized Upper Model*, referred to hereafter as GUM [18]. The GUM linguistic ontology provided a broad basis for driving natural language generation, but did not treat spatial language with any degree of sophistication. Therefore, in order to meet the requirements of a linguistically-motivated spatial organization as set out in this section, we have developed GUM further in three main respects:

- First, we now work with an explicit orientation to the state of the art in formal ontological engineering, observing modeling strategies from formal ontology [66,67,146] and enforcing criteria for ontological correctness from the OntoClean methodology for ontology evaluation [68].
- Second, we specify the ontology so that it is usable for both generation and analysis, thereby providing a common source and target semantic representation suitable for dialogue systems; this also provides a formal and logical representation which can be used together with automatic consistency checking and reasoning.
- And third, drawing strictly on linguistic evidence, we have combined both the results of the existing literature of empirical analyses of spatial language and conclusions from our own empirical studies of situated spatial discourse in order to produce a highly detailed extension of GUM for the spatial domain.

³ It is sometimes suggested that lexicalizations, particularly basic level lexicalizations, are revealing of deep underlying categories, but the extent to which this is true is debatable. We consider grammatical organization as necessarily a more robust indication of semantic import precisely because it needs to generalize across both situations and individual types of entities; lexical organization, in contrast, has the task of being specific, of *not* generalizing. We discuss this further in Section 3 on method below.

2.4. Conclusion

We have argued that folding together contextualization and linguistic semantics in the area of spatial semantics leaves an over-complex impression of the mechanisms involved. We need instead to isolate those linguistically-motivated semantic distinctions that provide a basis for contextualization and which themselves generalize across contexts of use. In particular, we need to provide an as exhaustive as possible characterization of the distinct kinds of spatial configurations and modes constructed by a language and to formalize how these are deployed in the construction of the shallow semantics of sentences as a whole. In Section 4 below we set out the categories of the linguistic ontology in detail; but first, in the following section, we prepare the ground for this by illustrating our methodology for designing a linguistically-motivated ontology such as the Upper Model. This will also let us introduce some of the top-level concepts of the ontology that we require below.

3. The development of a linguistic ontology of space

We emphasized in the previous section the importance we attribute to grammatical evidence. Focusing on the closed-class, or grammatical, end of the spectrum of linguistic phenomena allows stronger claims to be made concerning the coverage of the resulting semantic framework and its adequacy for explaining linguistic flexibility. Attempting a grammatical characterization for the spatial area is also of considerable interest for general linguistic reasons; Talmy claims, for example, that the grammatical category of closed-class forms in its *entirety* is actually limited to a small range of concepts coming from the spatial domain [154, p. 177]. In contrast, accounts starting from open-class items exhibit a tendency to slip over from accounts of language to accounts more appropriately considered as domain-specific extensions of non-linguistic ontologies. This is precisely the function of open-class items: to provide structure for all the various individual domains of experience at issue for a culture. Grammatical semantics, on the other hand, is there to provide a generalized foundation re-usable across domains [70].

Approaches that are not anchored in linguistics are naturally drawn to the lexical approach because it is the open-class items that are easiest to find without linguistic analysis. This also characterizes many computational and formal approaches to spatial language, together with a restriction of attention to a rather small subset of the linguistic phenomena that actually contribute to the linguistic construction of space. This is not adequate for building a linguistic semantics because its methodology is limited with respect both to the range of semantic distinctions that can be made visible and to the range of linguistic data that can be drawn upon as evidence. As pointed out with particular force by Levinson, spatial information is in fact distributed across *at least* determiners, adjectives, adpositions, relational nominals, adverbial nominals, grammatical case, locative verbs, verbal affixes, clitics and combinations of all of these [110, p. 99]; drawing on this range of evidence provides a much broader foundation for identifying what language itself is doing with space. It also, although we will not be able to address this here, prepares the ground effectively for accounts that apply *across* languages: as we shall see below, most of the distinctions drawn in our spatial extension to GUM can be found across languages, although their particular grammatical realizations may differ substantially.⁴

Our task here is to provide a level of semantic description that will support the process of contextualization. Such descriptions must distinguish the spatial situation picked out by an utterance from all other grammatically distinguishable situations without overcommitment: that is, the description captures formally just the degrees of flexibility, or underspecification, that the linguistic utterance itself leaves open. As mentioned above, we draw on several approaches that have been taken in the literature for developing a level of description for spatial language of this kind. In particular, we proceed in the directions proposed by Talmy [155], Bierwisch [26], Levinson [110] Halliday and Matthiessen [70], Eschenbach [53], Zwarts [169] and others, combining these within an overarching ontological framework anchored in formal ontology. This is organized in terms of appropriately motivated semantic types within the Generalized Upper Model linguistic ontology.

The move from grammatical evidence to a linguistic ontology can be complex and it is not feasible in this paper to show for each category just which range of linguistic phenomena is drawn upon; Tenbrink [157], for example, provides considerably more detail. It is, as a consequence, easy to lose sight of our fundamental commitment that no semantic distinctions be introduced unless we can specify their linguistic consequences. To show this process at work, however, we present two examples of linguistic analysis of this kind, starting from the grammatical constructions involved and relating these to the semantic configurations entailed. These configurations also provide our first views of the ontological organization of the Generalized Upper Model spatial extension.

⁴ This does *not* claim that all languages share all features; there are significant differences which represent important objects of study in their own right. There is also a need for considerably more study in the light of our argument below that full grammatical contexts need to be considered. Even traditional differences discussed in the literature, such as the contrast between English *in/on* and German and Dutch *in/an/auf*, *in/aan/op*, etc. or the well-known typological distinction between 'satellite-framed' and 'verb-framed' languages [153,145,154] reveal themselves to be less than clear-cut when naturally occurring usage data are considered.

3.1. Example 1: Basic configurations and the place of spatial relationships within these

The following is adapted from an original German example of a location description found in our own experimental studies on spatial language use (cf. Section 6):

(7) To the left of the computer is a USB drive.

All the usual potential ambiguities mentioned in Section 2 above apply to this utterance, although in this case we know precisely the spatial context and the dialogue development that led up to the utterance and so can reliably disambiguate. Without this context, however, **left of the computer** could mean “to the left from your point of view” just as well as “to the left from my point of view”, which can make a decisive difference if speaker and hearer do not share their view on the scene, or the speaker may ascribe an intrinsic left side to the computer—typically, the side that is on the left when interacting with it [73]—which is then independent of the current position of both speaker and hearer. As is overwhelmingly the case in natural interactions, the linguistic structure of the example chooses not to reflect these conceptual options, leaving it to the interaction and the hearer to resolve [157].

The utterance itself is an existential statement that ‘locates’ a *figure* with respect to a *ground*. The figure/ground distinction is widely accepted in spatial processing and perceptual psychology and is also regularly reflected grammatically in language. Talmy, for example, proposes that the figure is conceived of as a moving or movable entity which is more relevant in the present context, while the ground serves as a reference entity conceived of as more stationary [154]. This conceptual asymmetry is reflected as grammatical syndromes within linguistic structure—for example, by the grammatical distinctions of subject and (indirect) object, by syntactic distinctions between main and subordinate clauses, and by presenting some entities as presupposed and others as new. All languages have a range of ways of performing this particular placing of entities of different kinds in relation to one another, both semantically and syntactically, and we draw on this as evidence for distinguishing the particular ‘semantic configuration’ involved here from others. Grammatical evidence supporting this in the present case can be found in the kinds of *alternations* [109] possible (e.g., insertion of ‘there’ as grammatical subject), nonpossibility for progressive present tenses, of passive and many more that together make up distinctive syndromes of grammatical consequences (cf. [70]). It is these grammatical consequences that we use as signposts pointing to the semantic distinctions required of the Generalized Upper Model.

The syndrome of grammatical evidence clustered around the present example motivates the definition of a semantic entity called a Configuration. Configurations generally correlate with grammatical clauses or other grammatical units expressing events or relationships (e.g., nominalizations). Moreover, we can be more specific in that the particular type of clause in (7) corresponds to a particular subtype of Configuration, one that we call a SpatialLocating; we will see other subtypes of Configurations below. This type of semantic configuration has the function of relating some entity (the figure) to some specification of a place (the ground). The figure–ground asymmetry also motivates the assignment of distinct roles to the entities within the configuration: in the present example, the USB drive fills a locatum relation (the movable object to be located, or ‘referent’ [110, p. 39]), whereas the place (the ground) fills a placement relation.

The entity being placed and the place itself can also be distinguished from one another by virtue of distinct syndromes of grammatical behavior (e.g., most obviously, the former is a simple nominal phrase and the latter is a prepositional phrase). They are, accordingly, assigned to distinct linguistic-ontological categories within the Generalized Upper Model by drawing a corresponding high-level distinction between the concepts SimpleThing and Circumstance. This reflects the fundamental distinction between the *identity* of entities and the *location* of entities mentioned above (cf. [103,5]). In the example, the **USB drive** is a SimpleThing, and **to the left of the computer** is a further grammatically motivated subconcept of Circumstance, called a GeneralizedLocation. It is also worth noting here that, similarly to Zwarts [169], we adopt the position that individual spatial expressions (such as the prepositional phrase at issue here) receive their own individual spatial semantics independently of subsequent composition with events or objects—this is important for maintaining compositionality, to which we return below in Section 5.2.1.

The concept GeneralizedLocation provides a common semantics for any linguistic expression of a ‘place’—including prepositional phrases, adverbials, nominals and other spatially-relevant constituents when used in appropriate grammatical contexts. Place is therefore understood as a category constructed by the grammar of a language, in our case particularly English and German (but for other languages too: cf. the references to Talmy, Levinson and others above). In addition, although GeneralizedLocation *constructs* a place on the basis of some specified entity, or reference object, it does not do so directly but rather in terms of the distinct kinds of ‘modes of anchoring and locating’ that language(s) make available; these we group together under the single concept SpatialModality. Our linguistic ontology provides, again exclusively on the basis of grammatical evidence, a detailed characterization of the range of subtypes of SpatialModality and their interrelationships and constraints. We show this at length in the next section.

Formally, we may consider a GeneralizedLocation as a one-place predicate that ‘locates’ its argument. The precise placement of that argument is then *constrained* by the specification of a particular subclass of SpatialModality. This is captured in an expression of the form:

$$\lambda x. \text{LOC}(x, P_{\text{spatial-modality}}(R))$$

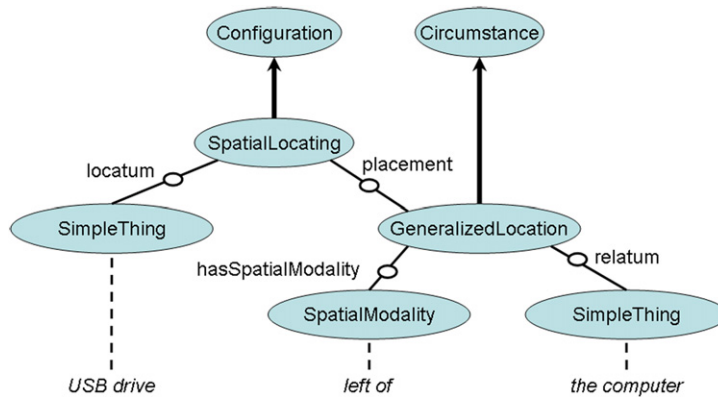


Fig. 3. GUM representation of example sentence (7) and its correspondence with the linguistic form.

where R is the reference object (e.g., the ‘desk’ in ‘in front of the desk’, the ‘tree’ in ‘in the tree’, etc.) and Loc a predicate of co-location.⁵ The expression says that some entity x co-locates with some ‘location’ that must be derived from the reference object R in the manner identified by the given spatial modality $P_{spatial-modality}$. This formalization is then similar to those proposed by Creary et al. [47] and Bierwisch and Wunderlich [25,164], both of whom argue for a complex compositional structure rather than superficially more straightforward two-place predicates. In our case, however, there will be no restriction to geometric interpretations of the reference objects for the reasons set out above.

This layer of interpretative indirection is the essential key to achieving an account adequate for the observed flexibility of linguistic spatial expression usage. Consider, for example, the kind of semantics that must be given for a prepositional phrase such as ‘at the town hall’ when used in the sentence:

(8) I am waiting at the town hall.

If we proceed directly to a contextualized interpretation of the spatial situation involved, we have a problem. Such a statement includes both waiting inside the town hall and outside it in the immediate vicinity as possibilities. No single ‘applicability field’ or selection of spatial region can be sensibly specified without considering the context of interpretation of the utterance. Moreover, as we saw in the examples of the previous section, this has nothing to do with a possible ‘relaxation’ or ‘stretching’ of the relevant applicability field: the range of application may turn out to be, as in the ‘between’ case illustrated in Fig. 2, totally disjoint. Employing our generic scheme for generalized locations, however, we can give the semantics of the prepositional phrase of (8) straightforwardly as:

$$\lambda x.Loc(x, P_{proximal}(R))$$

The concept proximal is one of the spatial modalities that we introduce in detail below. The expression as a whole means that a place is to be derived from the reference object, or relatum as we will now call it in line with general linguistic practice [19], subject to a constraint of Proximity. That is, the relatum, i.e., the town hall, is used to locate some entity via the mode of locating/anchoring of ‘being in a proximal relationship to’. Crucially, we take this already to be a *functional* characterization, since we cannot express metrically just what is proximal or not—this depends on the purpose of the interactants—nor over what duration the relation holds, whether the objects involved in a description are moving, and so on. All that the linguistic statement commits to is that the location is stably proximal ‘for current purposes’, whatever they might be. We do *not* at this stage, i.e., within the linguistic semantics, need to talk further of any kind of spatial region representing this proximity: linguistically, no further commitment is made beyond the claim of functional proximity and it is this that provides fixed semantic import for all constructions grouped under this particular linguistic-ontological category.⁶

Within our current example (7), the semantics is constructed similarly. **The computer** fills the relatum relation of its GeneralizedLocation (the more stationary object used for reference), while **to the left** is the spatial modality that is applied to the relatum in order to ‘create’ a place; we show the particular position of the spatial modality employed here, **to the left**, within the hierarchy of spatial modalities as a whole in Section 4 below. This abstract specification of a ‘place’ then fills the placement role of the SpatialLocating configuration and we arrive at the complete spatial semantics for the example utterance shown graphically in Fig. 3. In our discussions below, we will use standard Description Logic notation [9] for definitions rather than diagrammatic representations because, as set out in detail in Section 4, our linguistic ontology

⁵ We assume a standard compositional semantics, so for a preposition in a prepositional phrase, for example, the semantics of the phrase would in fact be the two-place predicate $\lambda R \lambda x.Loc(x, P_{spatial-modality}(R))$. The variable for the reference object is bound when the preposition and the prepositional object are combined within the prepositional phrase (cf. [53]). We omit this phrase-internal detail for the purposes of the current discussion.

⁶ The issue of what information is required where is an interesting one in its own right and needs to be subjected to empirical experimentation. Levelt [108], for example, discusses this with respect to perspective choice.

is in fact expressed within a description logic. This notation then provides a particularly succinct representation of the class–subclass relations and the availability of roles and possible constraints on those roles' fillers just as we require. The definition of *SpatialLocating* given so far (this will be refined further below) is then as shown in (9).

(9) *SpatialLocating* \equiv *Configuration* \sqcap \exists *locatum.SimpleThing* \sqcap \exists *placement.GeneralizedLocation*

Similarly, the concept *GeneralizedLocation* receives a definition of the form:

(10) *GeneralizedLocation* \equiv *Circumstance* \sqcap \exists *hasSpatialModality.SpatialModality* \sqcap \exists *relatum.SimpleThing*

These definitions form part of the linguistic ontology's *terminological* component (TBox), whereas the particular semantic representations constructed (indicated by the dashed lines in Fig. 3) are part of the *assertional* component (ABox) or instantial information. The non-diagrammatic notation we use for the instantial information will be introduced in Section 4.

Instantial representations, which can be derived automatically from the corresponding linguistic utterances (cf. Section 5.1), then need to be subjected to contextualization in order to reach a description of possible spatial situations. The semantic representation itself, however, only serves to constrain that interpretation by capturing the linguistic commitments entered into by the expression at hand. The area of spatial modalities in which **to the left** is located (*LateralProjections*: cf. Fig. 7), for example, only commits to a decomposition such as that involved in breaking some abstract space into half-planes formed with respect to a reference entity together with an imposed orientation (e.g., [53]). *LeftProjection* and *RightProjection* are then distinguished as two disjoint extensions of the theory of *LateralProjections*, anchoring further which of the half-planes is picked out and the relationships holding between the two. In each case we have a statement of fixed semantic import, expressed as a minimal abstract spatial 'theory'. The determination of the abstract spaces involved, the necessary reference entities, and their orientation then defines the range of possible interpretative uncertainty to be resolved during contextualization. The semantic representation makes the claim that all utterances sharing such a representation will face the same kinds of uncertainty regardless of how contextualization plays itself out in each particular case. We present examples of this in Section 5.2 below.

3.2. Example 2: Linguistically distinguished spatial situations

For our second example, we show in more detail how differences can be established between spatial modalities on the basis of the distinctive patterns revealed by grammatical syndromes. As a starting point, we take Talmy's suggestion that we can:

"...start with any closed-class spatial morpheme in any language, considering the full schema that it expresses and a spatial scene that it can apply to" [155].

Talmy illustrates this approach by drawing out the spatial schemas involved in the example:

(11) The board lay across the road.

This is another example of a relation between a figure and a ground. As before, we can consider this as correlating semantically with a *Configuration* involving a located entity (the board) and a generalized location constrained by a spatial modality. Talmy suggests that the greater part of the schematic structure in this case is conveyed by the closed-class term 'across' and, like other spatial closed-class terms, this expression represents a particular schematic spatial relationship between a (particular kind of) figure and a (particular kind of) ground. The question for us, and for the definition of the spatial modality component of the linguistic ontology, is just what the nature of that particular relationship is.

Talmy's method consists of setting out 'distinguishing scenarios' that differ minimally from the scenario associated with usage of the closed-class item *across* in English but which nevertheless call for the use of a different closed-class item or grammatical construction. This functions as an investigative probe for revealing those features deemed essential to the meaning of 'across'. If any of the features do not hold, or hold differently, then a different linguistic expression is required. Table 1 shows Talmy's results. Features (a) and (b) are taken as a starting point while features (c)–(i) are motivated by what happens linguistically when the corresponding features are violated in the distinguishing scenarios shown on the right. A similar kind of approach underlies many attempts to provide formal characterizations corresponding more closely to linguistic usage [8,54], including the development of specially tailored spatial calculi (e.g., [134]).

Talmy's methodology has been employed for a range of languages, revealing both similarities and differences across the semantic features those languages construct. But the tests are not exhaustive and can result in over-commitments: that is, the particular situations described in test sentences can lead to spatial distinctions being imported that are not linguistically motivated for the terms being investigated. A consideration of grammaticized distinctions is present in Talmy's account, but only implicitly. Talmy's description for 'across' is not, for example, an account just of the term 'across'—what is actually analyzed is the entire linguistic construction where 'across' appears, involving a spatial modality ('across'), a spatially-relevant activity ('lay') and the particular entities concerned (since both 'board' and 'road' bring with them spatial

Table 1

Features derived from appropriate or inappropriate spatial descriptions for 'across' according to Talmy [154]. F = Figure; G = Ground.

TARGET SENTENCE: <i>The board lay across the road.</i>		Distinguishing scenario
a.	F is spatially related to G	
b.	G is ribbonal	
c.	F is linear (and generally bounded at both ends)	vs. The wall siding lay over the road.
d.	The axes of F and G are roughly perpendicular	vs. The board lay along the road.
e.	F is parallel to the plane of G	vs. The board is sticking out of/into the road.
f.	F is adjacent to the plane of G	vs. The board lay (buried) in the road/ The board was suspended above the road.
g.	F's length is at least as great as G's width	vs. The baguette lay on the road.
h.	F touches both of G's edges	vs. The board lay over one edge of the road.
i.	The axis of F is horizontal (the plane of G is typically, but not necessarily, horizontal)/The spear hung across the wall.	vs. The spear hung up and down on the wall.

commitments). The description and motivation that Talmy offers of this method is then still, from our perspective, centred too narrowly on individual items rather than on what the grammatical organization and corresponding constructions as a whole commit to.

To show this more clearly we begin by noting that the semantics of 'across' as characterized by Talmy presupposes several features that are only partially due to the contribution of 'across'. As a case in point, Talmy proposes that 'across' is only appropriate due to the 'ribbonal' nature of the entity playing the role of ground: here, the road. But Talmy himself provides a further example in which the ground is not 'ribbonal' in the described sense (case (i)):

(12) The spear hung across the wall.

and, in fact, it is easy to find further counterexamples, such as "the bridge across the lake" which does not presuppose the form of the lake to be ribbonal. It seems more likely that the association of being 'ribbonal' is due to our world knowledge of those entities with which 'across' co-occurs, and so conceptualized information is again confused with the linguistic semantic contribution.

Within the Generalized Upper Model, we characterize the spatial semantics that appears to be involved in usages of 'across' in the intended sense by defining a particular spatial modality, that of *PathRepresentingInternal*. This modality constructs a 'place' by construing some path that is internal to another entity, given as the spatial modality's relatum, in the sense of 'lying within its borders'. Moreover, it is sufficient here to lie within the borders when the path is projected into the plane of the relatum rather than being strictly inside in terms of three-dimensional containment. The relatum then needs to be seen as (at least) 2-dimensional since 1-dimensional entities do not give rise to insides.

For this and similar reasons, GUM is also forced to draw a high level distinction between points (0 D) vs. lines and planes (1 to 2 D). This plays a role for the categorization of several spatial modalities, including both the present example and spatial modalities such as *Distribution*, which implies a complex (at least 1-dimensional) relatum. This latter modality is used in some configurations involving 'across' explicitly excluded from consideration in Talmy's example, as in the distributed meaning of 'across' in expressions such as "all across the country". In the present case, the entailed 2-dimensionality is consistent with a linguistic construal of the road as 2-dimensional.⁷ Putting this together, Talmy's finding that the figure needs to be one-dimensional is captured by the notion of 'path', while the ground needing to be planar is captured by the spatial modality being internal.

This contrasts with what happens when we take other spatial modalities: for example that corresponding with 'around' in phrases such as "the road around the lake". 'Around' is similar to 'across' in that some relatum is used to construct a 'path-like' place that locates the figure; however, it is also different in that the relationship is not necessarily *internal* to the relatum, although it might be, as seen in "He swam around the lake". As a consequence, we can only state that 'around' is *PathRepresenting* but the precise subcategory of *PathRepresenting* relevant for a particular usage is a pragmatic discourse inference made in connection with the relatum *and* the activity. Many prepositions (at least in English) under-commit in this way. These phenomena are captured by defining a more general category of *PathRepresenting* that is divided into two more specific categories, *PathRepresentingInternal* and *PathRepresentingExternal*; i.e.:

$$\text{PathRepresenting} \equiv \text{PathRepresentingInternal} \sqcup \text{PathRepresentingExternal}$$

Returning to our example, Talmy also suggests that some notion of *adjacency* must be involved since the board can be neither within the road nor somewhere above it. But this again involves several over-commitments. The observation is more appropriately accounted for by representing 'across' as committing only to the two-dimensional (topological) relationship (i.e., projecting the path into the plane of the relatum). It is then simply not possible to infer any internal relationship in the third dimension. As seen by "the bridge across the lake", the term 'across' as such can easily be used while violating

⁷ More precisely, this means that the sentence introduces a discourse reference object into the unfolding semantics of the text/dialogue that is construed semantically as 2-dimensional and which is used to refer to and classify some entity as a 'road'.

adjacency and so adjacency cannot be assumed to hold. This is also brought out clearly by the many examples in which ‘across’ and ‘over’, where adjacency is clearly not entailed, are interchangeable in context:

- (13) a. The shop is over/across the street.
 b. The bridge goes over/across the street.
 c. The man walks over/across the street.

In the present example, the adjacency component is actually contributed by the meaning of ‘lay’ rather than ‘across’, again showing that it is necessary to be sensitive to the entire grammatical construction when considering interpretations.

Talmy’s characterization of ‘across’ in terms of *boundaries* is also problematic: it is claimed that the figure needs to touch both edges of the ground. However, unlike formal geometric or topological considerations of space and regions, language does not rely on boundaries in this strict way. Instead, what is linguistically relevant here is the notion of covering the extension of the ground *functionally*: the board covers the full width of the road *for some purpose*. This generally remains the case even if it is somewhat shorter or considerably longer than the road’s extension. As soon as this sentence is set into a functional context, the exact topological relationship becomes irrelevant: a person could use the board (lying across the road) as a bridge over the muddy road even if a considerable piece was missing; this would not necessitate an expression like “the board lay over one edge of the road” (suggested by Talmy under case (h)) unless it became functionally relevant in the given context. Similarly, bridges are regularly built across rivers without touching edges or boundaries. Thus, touching the edges as such cannot be crucial to the linguistic semantics. Instead, the notion of *functionally covering the full extent* is the linguistically relevant characterization that needs to be captured as part of the shallow semantics of space. The concept *PathRepresentingInternal* represents this since it construes its path as being limited (functionally) by some entity’s boundaries. In contrast, realizations of *PathRepresentingExternal* do not necessarily involve full coverage, since paths may follow the outward shape of an entity for a limited extent.

‘Across’ also appears, as Talmy points out, to convey a particular relationship between the shape of the ground and the trajectory of the figure that Talmy characterizes as *perpendicularity concerning the axes* (case (d)). This must also be seen functionally since it is largely inferable from what it means to get ‘across’ an entity at all: i.e., a connection is made between one side of the relatum and another. The notion of perpendicularity at issue here is again therefore based on functional concerns rather than geometry. Some items, for example ‘through’, which can also serve as a realization of *PathRepresentingInternal* (since some ‘path-like’ place is created within the specified relatum), do not relate to perpendicularity (as in ‘he walked through the crowd’); other items, for example ‘over’, may be used in both perpendicular and *non-perpendicular* contexts. For example: in “the bridge leads over the street”, the bridge (as figure) is functionally perpendicular to the street (as ground), whereas in “walking over the bridge”, the trajectory of the path (as figure) is topologically parallel to the bridge’s linear extension (as ground).⁸

Many of the spatial notions that we can identify here are also related to construals of specifically constrained topological shapes for the ground. In the case of ‘across’, this concerns functional perpendicularity (as a notion specific to ‘crossing’) and coverage of the full extent (as a notion associated with the concept *PathRepresentingInternal*, encompassing several lexical realizations); Talmy’s *parallelness in the plane* is also captured in this way, since adhering to the ground’s topological contour is conceptualized as taking place in the plane. These properties provide further evidence for class–subclass relationships among the spatial modalities. In particular, the Generalized Upper Model makes the concept *PathRepresenting* a subconcept of the modality *ShapeCommitting*, indicating that the referent object (i.e., relatum or ground) has a determining effect for the shape of the ‘place’ that is constructed:

PathRepresenting \sqsubseteq *ShapeCommitting*

This dependence can vary from a quite specific determination (as in ‘following a path’ where the shape of the path is entirely responsible for the shape of the place constructed) to more generic determinations, such as that involved in getting from one side of the relatum to another side (as in ‘across the lake’). In the latter case, the shape is only constrained to be topologically bounded by ‘opposed’ sides that can be crossed between.

3.3. Conclusion

In this section, we have set out how we employ grammatical evidence in order both to motivate particular categories within our linguistic ontology and to articulate further the relationships and dependencies between these categories. We believe that the kind of approach proposed by Talmy brings out a rich collection of semantic features for consideration within an adequate semantics, but it is nevertheless striking that the restricted range of grammatical evidence that he draws upon reduces the accuracy and generality of the claims that can be made. The more semantic constraints we can motivate directly from patterns of linguistic usage, the more accurate our semantic formulation can become and the better we will be able to constrain interpretations of spatial terms according to the discourse purpose and the spatial situation.

⁸ Note, that there are topologically parallel relations that, when put into a real context, will not comply with a “walking over the bridge” scenario. Here, however, the examples only map from a linguistic description to certain topological constraints, not vice versa.

We have argued that this requires placing spatial terms in grammatically fleshed out contexts in order to have fullest access to the grammatical evidence available. A prepositional phrase such as ‘across the road’ cannot be considered in isolation because alone it represents an overly impoverished source of evidence. Such abbreviated starting points invite interpreters to ‘guess’ intended interpretations, which is a process subject to such broad variation as to be inherently problematic; this becomes even worse if we consider, for example, individual prepositions. Bringing in grammatical sources of evidence as we propose so as to consider entire constructional contexts (e.g., the clauses through which particular usages are expressed) allows us to go further than looking at particular lexicalizations while still constraining the range of interpretations that are considered eligible. This is seen as an essential pre-condition for bringing out the similarities and differences necessary for an accurate semantics.

Although space precludes giving the same level of detail for each category we discuss in this paper, in the development of the Generalized Upper Model linguistic ontology to which we now turn, we always base distinctions on as broad a range of grammatically complete examples as possible. Moreover, since we consider it unlikely that arbitrarily constructed examples provide sufficient variation to build upon, this also moves our method further towards empirical, corpus-based studies. This is brought out subsequently as one of the main criteria for evaluation we employ in Section 6 below.

4. The linguistically-motivated ontological structure of spatial categories and relations

We have now motivated a linguistic ontology for mediating between linguistic form and contextualized interpretation (Section 2) and sketched in what way grammatical evidence is taken as the principal guide to that linguistic ontology’s content and organization (Section 3). In this section we turn to the main results, the linguistic ontology itself. In particular, we set out in detail the spatial extension we have developed for the Generalized Upper Model linguistic ontology, showing how all of the categories discussed so far fit together within a comprehensive account of the linguistic semantics of space.

The spatial extension of GUM is a formal theory axiomatized in description logic (DL [9]), a decidable fragment of first-order logic (FOL). Using DL provides a structural backbone based on the class–subclass relationship (the ‘signature’ of the theory) and the usual description logic possibilities for defining roles between classes and constraints on those roles’ fillers. GUM’s signature therefore contains categories (unary predicates), also called classes or concepts, and relations (binary predicates); specific kinds of relationships are defined as relations between classes, which are themselves organized within a relation hierarchy.⁹ The GUM ontology itself is formulated in the Web Ontology Language OWL 2.0¹⁰ and is freely available (licensed under creative commons).¹¹ We employ both DL and OWL primarily because this guarantees the broad technical applicability of the ontology in diverse contexts of use, including natural language processing tools and dialogue systems, where DL, and now increasingly OWL, are established standards. We do not claim that the semantics of spatial language makes particular demands on expressivity that allow that semantics to be situated within one particular class of description logic rather than another, although this would be an interesting path to explore. Moreover, as the spatial extension of GUM is intended for use within running dialogue systems, our aim has been to keep the ontology within a decidable logic as far as possible. The classes making up the class–subclass backbone and the roles defined over these classes are motivated throughout by distinguishable grammatical syndromes of the kind suggested in the previous sections. Standing ‘behind’ the concepts of the spatial extension backbone, we envisage modular ‘spatial theories’, represented as formal specifications, that capture the minimal *semantic* commitments entered into by each class (cf. [19,80] and Section 5.2); in this section, however, we focus on the backbone.

The most general distinctions within the GUM concept hierarchy partition the top node, GUMThing, into three basic subcategories: Configuration, Element and MultiConfiguration. These categories represent three distinct scales of semantic complexity for entities, related in various ways to one another as we shall see below. A Configuration in GUM is, as described informally above, the semantic correlate of some activity or state of affairs, i.e., some representation of experience [18]. The semantics of simple clauses therefore corresponds, in general, to Configurations and the description offered by GUM is most closely related to event-based representations in linguistics drawing on a Neo-Davidsonian linguistic semantics [49]. Combinations of clauses correspond to MultiConfigurations—these are specified as semantic dependency structures holding over configurations.

Subtypes of Configurations define distinctive sets of relations which typically take Elements as their fillers. These relations then capture the different ways in which Elements can participate in events, activities, or states. Exactly one of those elements has to stand in the relation *processInConfiguration* and is restricted to be a particular subtype of Element called a Process. Processes are most commonly expressed linguistically by verbal groups [71]. Further participants in the state of affairs are related in configuration-specific ways to the Configuration as a whole and define those semantic elements which are necessarily entailed by the kind of configuration occurring: e.g., a ‘running’ state of affairs can only exist together with something that ‘runs’, etc. Additional information concerning when, where, how, or under which conditions some state of affairs occurs is provided by a further subtype of Element called Circumstance, related to the configuration by a small

⁹ Throughout this paper we use ‘category’, ‘concept’, ‘class’, ‘type’, etc. as synonyms, reflecting the fact that several different traditions are being combined in the discussion and DLs do not distinguish these formally in any case; similarly for ‘relation’, ‘role’, ‘feature’, ‘property’, ‘attribute’, etc.

¹⁰ <http://www.w3.org/TR/owl2-syntax>.

¹¹ <http://www.ontospace.uni-bremen.de/ontology/stable/GUM-3-space.owl>.

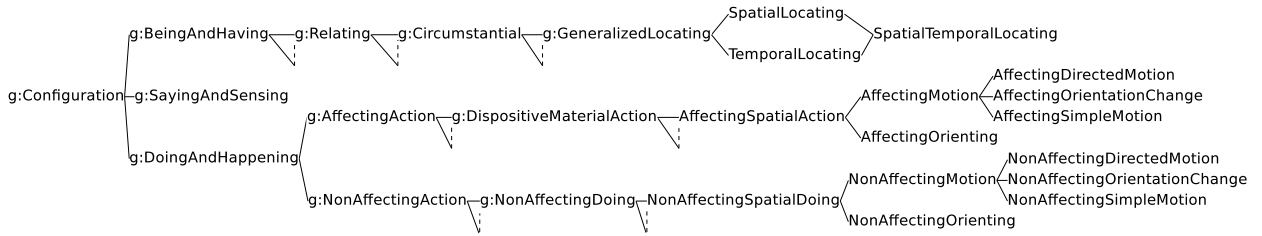


Fig. 4. Taxonomic hierarchy of spatial configurations in the GUM concept hierarchy. Dotted lines indicate further subconcepts, the namespace “g” indicates that the concept is defined in the core GUM ontology, while terms without a namespace are defined in the spatial extension.

number of configuration-independent relations. The basic definition of a Configuration, given again in standard Description Logic notation [9], is then:

$$(14) \text{ Configuration} \equiv \text{GUMThing} \sqcap \exists \text{processInConfiguration.Process}$$

i.e., any configuration is a subclass of GUMThing and necessarily has a role processInConfiguration filled by a Process. In addition, subclasses of configurations define further participantInConfiguration and circumstanceInConfiguration relations, filled by SimpleThings and Circumstances respectively; we omit a formal specification here. These distinctions are, as always, demanded by distinctions made linguistically: semantic participants often correlate with obligatory grammatical elements, and circumstances with optional ones (cf. [71]). Both of them, however, can be expressed by nominal groups, adverbial groups and prepositional phrases.

Within English, German, and many other languages there is strong evidence for several immediate subcategories of Configuration (cf. [70]). Within GUM we adopt the following three disjoint Configuration subconcepts: BeingAndHaving, DoingAndHappening, and SayingAndSensing. BeingAndHaving configurations describe statements of existence and static relations between participants, DoingAndHappening configurations describe intentional actions and non-intentional happenings, and SayingAndSensing configurations represent cognitive and verbal acts. As we are addressing the linguistic-ontological structure of spatial language here, we will focus on those configurations, elements, and relations involved in the representation of spatial semantics. The first few steps in the configuration sub-hierarchy leading down to spatially-relevant categories are shown in Fig. 4. Again, these broad configuration classes are all distinguished by differing grammatical behavior. As noted above, we will not always have the space to present this grammatical evidence explicitly and so it is important to recall that *no distinctions are introduced into GUM without linguistic grounding* regardless of whether we explicitly show that grounding here. Moreover, although we sometimes employ simplified constructed examples to illustrate concepts, our main evidence is naturally occurring corpus or experimental data. All of the distinctions are therefore motivated from linguistic phenomena as set out in the discussion of the previous sections.

We have already seen one particular case of a Configuration that involved spatial semantics: both the main examples of Section 3, i.e., (7) and (11), were assigned to the semantic class SpatialLocating. The position of this class in relation to other GUM configurations can be seen in Fig. 4, where we see several further subconcepts lying between SpatialLocating and Configuration; we describe those that are spatially relevant below. Several other branches in the Configuration hierarchy also contain spatially-relevant configurations; alongside SpatialLocating two other configurations are particularly relevant:

- (i) NonAffectingSpatialDoing: dynamic spatial configurations, in which an entity is changing its relative position (such as motion or orientation);
- (ii) AffectingSpatialDoing: dynamic spatial configurations, in which an entity affects the relative position of another entity (such as motion or orientation).

These are both subconcepts of the main configuration type DoingAndHappening and so correlate with very different grammatical phenomena to those of the concept SpatialLocating, a subconcept of the main class BeingAndHaving.

The inter-connections between these distinct types of configurations and the many subclasses of spatial modalities will become quite detailed as we proceed. Fig. 5 therefore gives a graphic representation of the general structure of types of classes and relations in GUM relevant for specifying the semantics of spatial language as a road map of what is to come. In this style of representation, adopted from the documentation of the DOLCE foundational ontology [119], classes are represented as boxes and class-subclass relationships are shown by spatial inclusion (i.e., boxes within boxes). The relations between classes are indicated by directed gray arrows. Taking the lowest two relations in the center of the figure, for example, we can read off the diagram that the relation actor is a relation between a SimpleThing (which is a subclass of Element) and DoingAndHappening (which is a subclass of Configuration), while actee is a relation between SimpleThing and AffectingAction (which is one of the two main subclasses of DoingAndHappening). We refer back to this diagram at many points in the discussion below.

Our overview of the GUM linguistic ontology is structured as follows. First we provide a more complete description of the SpatialLocating configurations that we have seen and used in our examples above (Section 4.1). Then we look at non-

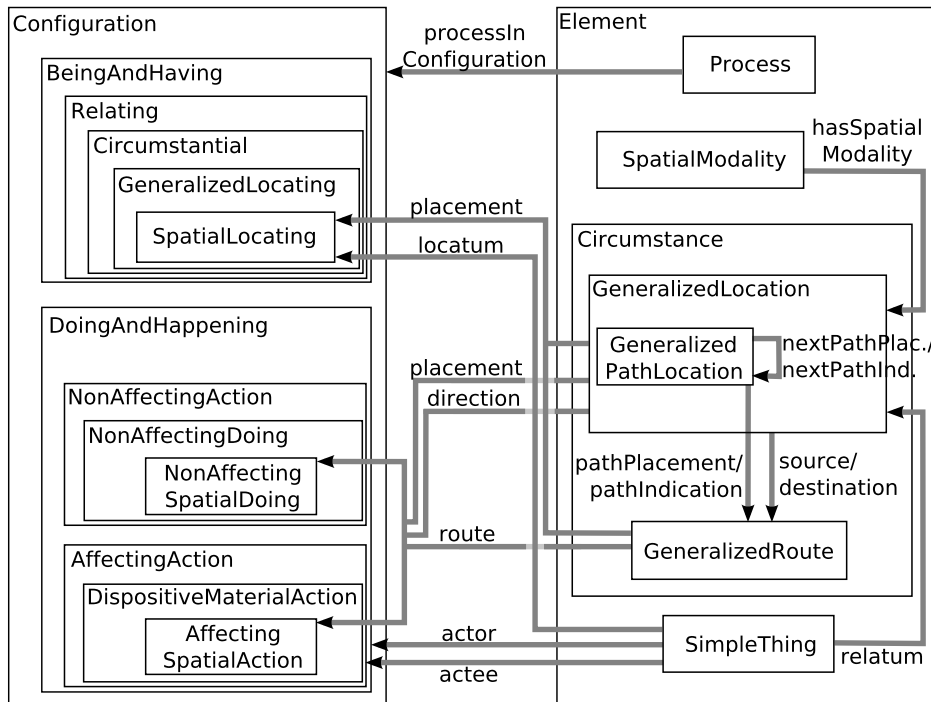


Fig. 5. Overview of GUM's spatial concepts and their relations.

affecting and affecting dynamic configurations (Sections 4.2 and 4.3 respectively). Following this, we set out the distinct kinds of spatial elements used in these configurations, namely: GeneralizedLocation and an extended kind of 'place', called a GeneralizedRoute (Section 4.4), then the SpatialModalities themselves (Section 4.5), and finally those SimpleThings that exhibit particularly spatial aspects (Section 4.6).

4.1. Static spatial configurations and their elements

In English, German, and many other languages, there is a broad class of clausal constructions distinguished by distinctive syndromes of grammatical properties that may be termed 'relational' (cf. [71]). These bring together entities and attributes of those entities. In regular, non-spatial relational expressions, such as "The book is old", attributes (e.g., "old") are assigned to entities (e.g., "book") by the GUM configuration of Relating, a subconcept of BeingAndHaving. Within the Relating configuration and all its subtypes, the entity and the attributes are distinguished by differing relations: the GUM relation hierarchy specifies the relation domain for the entity and the relation attribute for the attribute.

Proceeding downward in the configuration hierarchy below Relating (cf. Fig. 4) is the concept Circumstantial; configurations of this type relate an entity to some attribute of cause, reason, purpose, or other circumstance-like attribute, including those involving space and time. For these latter circumstances, we have a further particular subtype of Circumstantial, a GeneralizedLocating. For configurations of this type, the required relation domain is further specified in the relation hierarchy to be the subrelation locatum, corresponding as we have seen above to the 'located entity'.

GeneralizedLocatings cover both spatial and temporal locatings. A configuration defining only a spatial relation is then a SpatialLocating, while a configuration defining only a temporal relation is a TemporalLocating. Focusing on the SpatialLocating case, the inherited attribute relation is further specified to be a placement relation; this allows the configuration to specify the place where the entity is being positioned; we showed in Fig. 3 above how this configuration is structured internally. The filler of the placement is in fact restricted to be one of two subtypes of Element, however: either it is a GeneralizedLocation, as seen in our first examples, or it is a GeneralizedRoute. We describe GeneralizedRoutes in more detail below; for the present we note that it is equally possible to specify a place in terms of a single location ("the dog is in the park") and as an 'extended' place ("the road runs from Bremen to Hamburg"). The grammatical behavior of the clause construction in this latter case is far closer to the first example than it is to a motion event and so we capture this within GUM by allocating them to the same configuration, i.e., a SpatialLocating; the internal behavior of places and routes is, however, different and so they are distinguished within GUM. GeneralizedRoutes may specify start, intermediate, and end points of the 'route' [50] and that route may then be specified by both actions and attributions of spatial extent. Jackendoff [86] and others propose that humans distinguish conceptually as well as linguistically between *places* and *paths* and the GeneralizedLocation vs. GeneralizedRoute distinction reflects this. We gave a preliminary definition of SpatialLocating in (9) in Section 3 above; now we refine the definition to fit with the GUM hierarchy in full:

- (15) SpatialLocating \equiv Circumstantial \sqcap \exists locatum.SimpleThing
 \sqcap \exists placement.(GeneralizedLocation \sqcup GeneralizedRoute)

The definition we gave for GeneralizedLocation (10) is unchanged; we repeat it here for ease of reference:

$$\text{GeneralizedLocation} \equiv \text{Circumstance} \sqcap \exists \text{hasSpatialModality.SpatialModality} \sqcap \exists \text{relatum.SimpleThing}$$

As observed above, although the relatum relation is necessarily present, it is often left unspecified in natural discourse [157] and so must be filled in during contextualization.

The GeneralizedLocation binds together a relatum and a spatial relationship within a single structured entity that may stand in a placement relation within a spatial configuration. An encapsulation of this kind has been argued better to support compositionality (cf., e.g., [47,168]) and also allows various combinations of multiple locations to be expressed within single configurations, as in the example “The plant is in the corner, by the window, next to the chair”, in which one SpatialLocating defines three placements. This becomes even more important when, as described below, placements are modified by expressing spatial perspectives, spatial accessibility, as well as extensions or enhancements of the spatial relation. For example, the utterance “The plant is to the front left of the chair, right here in the corner” combines two relations (front and left) with respect to one relatum (the chair), while “in the corner” is enhanced with possible access information “right here”. Modifications contained within single encapsulated ‘places’ are easier to track when they are re-used across spatial placements in a dialogue [47]. This is a further motivation for defining GeneralizedLocations so that they retain their structure independently of the configurations in which they are used. Thus we find the ‘same’ GeneralizedLocation at work in the configurations corresponding to “he goes to the right of the chair” (dynamic spatial configuration) and “he is standing to the right of the chair” (static spatial configuration)—the difference being captured, as we will see below, in terms of two distinct modes of participation in their configurations: i.e., direction (cf. (18)) and placement (cf. (15)) respectively. We can also now combine these definitions and show how particular spatial statements are represented. For this we adopt a further standard notation used for description logic assertional, or instance information, that of the Manchester syntax for OWL 2.0.¹² Using the classes and relations defined within GUM, the clause “The USB drive is on the table” receives the semantics shown in (16), consisting of two related instances, or ‘individuals’.

- | | |
|--|--|
| <p>(16) Individual: <i>SL-of-USB</i></p> <p>Types: SpatialLocating</p> <p>Facts: locatum <i>The USB drive</i>
 placement <i>GL-of-USB</i></p> | <p>Individual: <i>GL-of-USB</i></p> <p>Types: GeneralizedLocation</p> <p>Facts: relatum <i>Table</i>
 hasSpatialModality <i>Support</i></p> |
|--|--|

All instance names (shown in slanted font) appearing here and in similar expressions used in this paper are to be read strictly as placeholders or semantic variables—that is, they do not indicate any hidden lexical or syntactic information; we will also regularly omit non-spatial information. Although we will return to this point again in Section 5.1 below, it is nevertheless worth emphasizing here that this kind of assertional specification is generally to be derived automatically using computational analysis grammars, drawing on the close connection between linguistic forms and semantic specification guaranteed by our reliance on grammatical syndromes when constructing the ontology in the first place.¹³

Finally, spatial location descriptions can also occur *together* with temporal specifications. This may happen in two ways: either both pieces of information are linguistically presented as equal, or the temporal information is added to the spatial location information. As an example of the former case, the sentence “The meeting is in the city center at eight o’clock” can be split linguistically into spatial and temporal components, as in “The meeting is in the city center and the meeting is at eight o’clock”. Therefore, the spatial and the temporal entities are of equal status and the sentence is a realization of the configuration SpatialTemporalLocating (cf. Fig. 4). An example of the latter case is: “He was in the city center at eight o’clock”; this sentence cannot be split into “*He was in the city center and he was at eight o’clock”, showing grammatically that the temporal component depends on, or further modifies, the attribution of spatial information. This sentence is then a straightforward realization of SpatialLocating with additional circumstantial temporal information.

The two sentences also differ with respect to the kind of entity being talked about, i.e., that which fills the locatum relation inherited from GeneralizedLocating. In a SpatialTemporalLocating, the locatum must be an event; whereas a SpatialLocating does not require any specific kind of locatum, i.e., it may be a person, an object, an event, etc. This distinction reflects the syndrome that events are linguistically construed as spatiotemporal entities; they can be referred to either in terms of their spatial or in terms of their temporal attributes. Persons and objects also exist in time and space, but English and German do not construe them temporally: “*He was on Monday.”

¹² <http://www.w3.org/2007/OWL/wiki/ManchesterSyntax>.

¹³ Note that this property also has consequences for evaluation. For example, when our ontology assigns a linguistic structure to an ontological partition, one can predict corresponding grammatical alternations for this structure, particular semantic entailments, as well as possibilities for contextualization, all of which can be tested.

4.2. Non-affecting dynamic spatial configurations and their elements

With actions or events such as “I ran” or “the ball rolled” either the state of affairs is construed linguistically as an actor intentionally performing an action or as an unintentional event occurring. Such configurations are often expressed by intransitive verbs or, if transitive, the object is not affected or created by the action or event but serves instead to restrict or further define the type of action in question—as in “I play tennis” [70]. Within GUM, such configurations are specified as *NonAffectingActions*, a subconcept of *DoingAndHappening* (cf. Fig. 4). These require one additional obligatory participant, the actor, for the entity who performs the action or who initiates the event; the respective dependencies can be seen in Fig. 5. *NonAffectingActions* are themselves subdivided into configurations that represent either actions or events. Configurations representing actions have one spatially relevant subconcept, *NonAffectingSpatialDoing*, in which an actor performs a spatial action. Such actions may be associated with spatial direction terms that express a particular orientation or movement in a specific direction [160]. Accordingly, there are two subclasses of this concept:

- (i) *NonAffectingOrienting*, which defines a spatially oriented relation of the actor, for instance “We are facing the table”. Here we have a *GeneralizedLocation* as usual, but in this case it fills the configuration-specific relation *orientationDirection*.
- (ii) *NonAffectingMotion*, which defines a spatial movement of the actor, for instance “He is walking towards the table”.

Movements are the more complex of the two and can be expressed by directed motions, re-orientations, or under-specified movements of the actor. The category *NonAffectingMotion* is accordingly divided into three subclasses:

- (17) $\text{NonAffectingMotion} \equiv \text{NonAffectingDirectedMotion} \sqcup \text{NonAffectingOrientationChange} \sqcup \text{NonAffectingSimpleMotion}$

In *NonAffectingDirectedMotions*, the direction of motion can be expressed in different ways. For example, if a direction is specified without a specific goal, the direction depends only on the actor, as in “He walks forward” or “He goes upward”. The *GeneralizedLocations* here fill a configuration-specific relation *motionDirection* and cannot freely define an additional relatum; it is the position of the actor which determines the interpretations possible for ‘forward’, ‘upward’, etc. Alternatively, if the direction is expressed with specific entities, as in ‘along the road’, the configuration defines the relation *route*. The definition of *NonAffectingDirectedMotion* is then as follows:

- (18) $\text{NonAffectingDirectedMotion} \equiv \text{NonAffectingMotion} \sqcap (\exists \text{direction.GeneralizedLocation} \sqcup \exists \text{route.GeneralizedRoute})$

GeneralizedRoutes can define the relations *source*, *pathPlacement*, *pathIndication*, and *destination* for their route components, and these in turn are filled by *GeneralizedLocations*; we will see the definition in detail below. The instantial semantics shown in (19) gives an example of a specification that defines the start and end point of a route in a *NonAffectingDirectedMotion*. This is the GUM specification corresponding to the expression “The deer ran from the hill to the stream.”

(19) **Individual:** *NADM*

Types: <i>NonAffectingDirectedMotion</i>	Individual: <i>GLsource</i>
Facts: actor <i>Deer</i>	Types: <i>GeneralizedLocation</i>
processInConfiguration <i>running</i>	Facts: relatum <i>Hill</i>
route <i>GenRoute</i>	hasSpatialModality <i>GeneralDirectionalDistancing</i>
Individual: <i>GenRoute</i>	Individual: <i>GLdest</i>
Types: <i>GeneralizedRoute</i>	Types: <i>GeneralizedLocation</i>
Facts: source <i>GLsource</i>	Facts: relatum <i>Stream</i>
destination <i>GLdest</i>	hasSpatialModality <i>GeneralDirectionalNearing</i>

Motions that do not involve a change of location but only a change of orientation are specified by the configuration *NonAffectingOrientationChange*. The direction of the orientation is defined by the relation *orientationDirection*, which can be expressed by a directional term, such as “The wheel rotates to the right”, or by a reference object together with a spatial relation, such as “He curves toward the jars”. *NonAffectingOrientationChanges* can also define sources and destinations, given by a route, in which case the direction can be expressed with several segments, such as “turn from the kitchen (90° to the left) to the office”. Here both the *GeneralizedLocations* “from the kitchen” and “to the office” are defined as the source and destination in a *GeneralizedRoute* since they can also indicate start and end points of a directed motion; intermediate places within a turn, however, appear to be rare. The resulting definition is given in (20).

- (20) $\text{NonAffectingOrientationChange} \equiv \text{NonAffectingMotion} \sqcap (\exists \text{orientationDirection.GeneralizedLocation} \sqcup \exists \text{orientationRoute.GeneralizedRoute})$

The third category, *NonAffectingMotion*, does not express an orientation or direction but a particular kind of spatial motion. Examples for this category are “He is wiggling up and down” or “He is dancing around”. The relation placement is then specified for the places expressing where the action takes place; e.g. “He is dancing around *in the street*”.

4.3. Affecting dynamic spatial configurations

In expressions of actions or events such as “They are assembling the engine” or “The drink was placed on the table before him”, an actor either creates or affects an entity. In GUM, such expressions are represented by the configuration *AffectingAction*, a further subconcept of *DoingAndHappening*. It inherits the relation *actor* for the actor and defines the additional relation *actee* for the affected entity (cf. Fig. 5); i.e.:

$$(21) \text{ AffectingAction} \equiv \text{DoingAndHappening} \sqcap \exists \text{actee. SimpleThing}$$

AffectingActions either create an entity that did not exist before the action, represented by the subconcept *CreativeMaterialAction*, or they affect (for instance change, manipulate, or destroy) an entity that is construed as existing, represented by the subconcept *DispositiveMaterialAction* (cf. [70]).

For the GUM spatial extension, a necessary subconcept of *DispositiveMaterialAction* is *AffectingSpatialAction*, a configuration that defines an action in which an entity is spatially affected by an actor. In general, the actor changes the location or orientation of the actee. Similarly to *NonAffectingSpatialDoing*, *AffectingSpatialAction* has two subconcepts, i.e., *AffectingOrienting* and *AffectingMotion* (see Fig. 4). Apart from defining the relation *actee*, their definitions correspond to those of non-affecting orienting and motion configurations respectively. This correspondence also holds for the subconcepts of *AffectingMotion*, yielding the corresponding subconcepts: *AffectingDirectedMotion*, *AffectingOrientationChange* and *AffectingSimpleMotion*. An example of an *AffectingDirectedMotion*, “He puts the USB drive on the table”, is given in (22); here, the “USB drive” is the actee spatially manipulated by the actor.

<p>(22) Individual: <i>ADM</i> Types: <i>AffectingDirectedMotion</i> Facts: actor <i>He</i> actee <i>USB drive</i> processInConfiguration <i>putting</i> route <i>GenRoute</i></p>	<p>Individual: <i>GenRoute</i> Types: <i>GeneralizedRoute</i> Facts: destination <i>GLdest</i> Individual: <i>GLdest</i> Types: <i>GeneralizedLocation</i> Facts: relatum <i>Table</i> hasSpatialModality <i>Support</i></p>
--	--

The actor in *AffectingSpatialActions* may also remain unspecified in the linguistic form. This often occurs in passive constructions, such as “The cupboard was moved to the wall”, or in ‘middle’ constructions, such as “The cupboard goes over there”.

4.4. Spatial elements I: GeneralizedLocation and GeneralizedRoute

Although static and dynamic descriptions clearly exhibit some fundamental conceptual differences [123], the way in which language construes static and dynamic situations shows many points of overlap. In English for example, it is not possible without further grammatical context to state whether ‘across the street’ is static or dynamic. Static contexts involve a figure and a ground construed linguistically as two currently immobile entities that are spatially related to one another in a particular way defined by the spatial modality. Dynamic contexts may involve the same spatial schema but in these cases the figure is not an immobile entity but rather the *trajectory* covered by some entity. There are also intermediate forms in which the entities involved are static, yet there is an entailed trajectory that licenses the spatial term, as in “The shop across the road”, which is motivated by the necessity of crossing the road in order to reach the shop.

In all these cases, it is the grammatical syndromes of the clauses as a whole that provide distinguishing information. GUM therefore represents, as mentioned above, elements that can occur in both static and dynamic configurations in the same way. Dynamic descriptions can, however, give more spatial information than static descriptions: for example, they can specify a path, a starting point or an end point of an action. Our formalization reflects this, first, by distinguishing *GeneralizedLocations* from *GeneralizedRoutes* as indicated above and, second, by providing more attributes for dynamic configurations than for static configurations.

As already shown, *GeneralizedLocations* describe relative positions of entities and fill a variety of relations, such as a placement of a static location, a direction of a directed movement, an orientationDirection of an orientation change motion, or a motionDirection in a directed motion. They are also used to specify parts of a route, resulting in the following definition for the category *GeneralizedRoute*:

$$(23) \text{ GeneralizedRoute} \equiv \text{Circumstance} \sqcap (=1 \text{ source.GeneralizedLocation} \sqcup =1 \text{ destination.GeneralizedLocation} \sqcup \geq 1 \text{ pathPlacement.GeneralizedPathLocation} \sqcup \geq 1 \text{ pathIndication.GeneralizedPathLocation})$$

source and destination specify the start or end point of the movement, while the relations *pathPlacement* and *pathIndication* specify intermediate points. These relations are each filled by a *GeneralizedPathLocation*, a further subconcept of *GeneralizedLocation* that adds in the possibility of building up chains of connected locations serving as route segments.

Placements and indications of paths are distinguished according to the way the path is described. A *pathPlacement* is defined when the relatum is linguistically construed in 1, 2, or 3 dimensions and locates the path itself; examples of path placements are “along the river”, “down the corridor” or “through the tunnel”. A *pathIndication* is defined when the relatum is linguistically construed in zero dimensions and gives a single point constraining where the path may lie; examples of path indications are “by the tree”, “past the fence”, etc. When a *GeneralizedRoute* is related to an orientation change configuration, however, neither *pathIndication* nor *pathPlacement* can be defined; this reflects the effects of the distinctions in the dimensionality of entities discussed in previous sections.

The instantial semantics shown in (24) gives an example of a more complex specification combining locations and generalized routes. This extends the example given as (19) above, showing the GUM specification corresponding to the expression “The deer ran from the hill down the road by the old tree to the stream.”¹⁴

(24) **Individual:** *NADM*

Types: *NonAffectingDirectedMotion*

Facts: actor *Deer*
processInConfiguration *running*
route *GenRoute*

Individual: *GenRoute*

Types: *GeneralizedRoute*

Facts: source *GLsource*
pathPlacement *GLpathP*
pathIndication *GLpathI*
destination *GLdest*

Individual: *GLsource*

Types: *GeneralizedLocation*

Facts: relatum *Hill*
hasSpatialModality *GeneralDirectionalDistancing*

Individual: *GLpathP*

Types: *GeneralizedPathLocation*

Facts: relatum *Road*
hasSpatialModality *PathRepresentingInternal*

Individual: *GLpathI*

Types: *GeneralizedPathLocation*

Facts: relatum *OldTree*
hasSpatialModality *Proximal*

Individual: *GLdest*

Types: *GeneralizedLocation*

Facts: relatum *Stream*
hasSpatialModality *GeneralDirectionalNearing*

Whereas the route parts in this example co-describe the route given, it is also possible to construct a route out of a sequence of route segments so that one segment is understood to *follow* another. The grammatical evidence for these distinctions is complex and builds on the differing ordering possibilities and preposition choices available for the distinct types. An example of such an expression is “along the river, over the bridge, past the fence and down the road”. Here, the whole path is represented by a *GeneralizedRoute*, specifying the first placement “along the river” by a *pathPlacement* filled by a *GeneralizedPathLocation*. This element then defines not only the spatial modality (“along”) and the relatum (“the river”) but also a further relation *nextPathPlacement*, which is itself filled by another *GeneralizedPathLocation* representing “over the bridge”. *GeneralizedPathLocations* not only inherit the relations *hasSpatialModality* and *relatum* from *GeneralizedLocation*, but also specify a successor relationship for positioning route segments within an ordered list of *pathPlacements* or *pathIndications*. Two successor relations are distinguished for this: *nextPathPlacement* and *nextPathIndication*, according to the type of route segment following.¹⁵

One last subconcept of *GeneralizedLocation* is *GeneralizedComplexLocation*. This element represents a location requiring a particular kind of ‘complex’ relatum for constructing a place. Such complex entities either consist of more than one element, i.e., they are linguistically expressed as plurals, for instance “He went between the objects”, or are themselves subconcepts of the GUM category *DecomposableObject*, for instance “He stood among the crowd”. *GeneralizedComplexLocations* can, as with all *GeneralizedLocations*, be used in both static or dynamic spatial configurations but select particular spatial modalities: in particular, subclasses of the spatial modality *Distribution* (see Section 4.5).

Finally, all *GeneralizedLocations* may additionally specify the following relations:

- accessibility, when some kind of specific access between the locatum or actor/actee and the *GeneralizedLocation* is expressed; this accessibility might be high or low, as in “I am *here*” vs. “I am *there*”.
- reciprocity, when a spatial reciprocal relationship between the locatum or actor/actee and the relatum holds; an example of such a relationship is “The cups are beside *each other*”.

¹⁴ ‘Down’, in ‘down the road’, is represented as a *PathRepresentingInternal*, which, as we saw in Section 3.2, is also the spatial modality used for ‘across’.

¹⁵ The grouping of places within a *GeneralizedRoute* corresponds closely to a semantic interpretation in which places are intersected; in contrast, the sequences of places created by *nextPathPlacement* or *nextPathIndication* correspond to a concatenation of paths as described by Zwarts [169]; we return to this in Section 5.2.

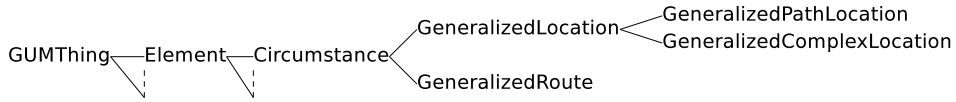


Fig. 6. Spatial circumstances in the GUM concept hierarchy.

- spatialPerspective, when a spatial perspective is expressed for a spatial configuration, e.g., “The desk is to the right from my point of view” or “Seen from above the helicopters move to the left”.

The distinct kinds of locations and routes defined in GUM are summarized in Fig. 6.

4.5. Spatial elements II: SpatialModality

As we have motivated and illustrated in previous sections, the core of the GUM treatment of spatial relations is provided by the internal organization and subconcepts of the concept SpatialModality. This is the part of the linguistic ontology that corresponds to the type of relationship being described in any linguistic spatial description, typically expressed grammatically by a spatial preposition, an adverb, an adjective, a part of the verb, or as entailed by the lexical semantics of the verb. SpatialModality fills the hasSpatialModality relation in GeneralizedLocations and specifies the type of spatial relationship to be construed between a locatum and a relatum in a static configuration, or between an actor/actee and a relatum in a dynamic configuration. An overview of the subconcepts of SpatialModality is shown in Fig. 7.

The most general distinction is whether the modality constructs distance between entities (SpatialDistanceModality), functional dependencies between entities (FunctionalSpatialModality), or positions between entities relative to each other (RelativeSpatialModality):

$$(25) \text{ SpatialModality} \equiv \text{SpatialDistanceModality} \sqcup \text{FunctionalSpatialModality} \sqcup \text{RelativeSpatialModality}$$

This disjunction is not exclusive; spatial modalities may fall under several of the main categories, as described below. Different spatial modalities can also be combined by *extension* (cf. [70])—for instance “left” extends “front” in the expression “it is to the front left of him”.

Spatial modalities can also be modified in order to give a more explicit description of the relationship between entities [168]. Such modifications are often used in order to describe the placement or the motion more precisely. The following modifications are specified in GUM:

- Qualitative or quantitative spatial information *enhances* spatial modalities, for instance “diagonally” enhances “left” in the expression “The TV is diagonally to the left”; in this case the SpatialModality defines the relation hasEnhancement, filled by QualitativeSpatialTemporal or QuantitativeSpatialTemporal entities.
- Spatial modalities can also be modified by *extreme positions*, which can be expressed with respect to a specific distance or axis. They are usually expressed by superlatives, such as “the leftmost” or “the furthest left” for extreme axis and “the closest” or “the farthest” for extreme distance information. For these, SpatialModality defines the relations extremePositionOnAxis or extremeDistancePosition.
- Angles and distances can be combined with spatial modalities, specified by the relations qualitativeAngleExtent, quantitativeAngleExtent for angle information and qualitativeDistanceExtent or quantitativeDistanceExtent for distance information. Angles and distances can either be expressed qualitatively or quantitatively and the respective relations are subsumed under qualitativeSpatialExtent and quantitativeSpatialExtent. Examples are “They turn around 180°” (quantitativeAngleExtent), “Turn a bit further to your right” (qualitativeAngleExtent), “He moves a bit further forward” (qualitativeDistanceExtent) and “They go three steps left” (quantitativeDistanceExtent).¹⁶

4.5.1. Spatial relationships about distance

A SpatialDistanceModality subsumes expressions about spatial distances between entities. Examples of this concept are “The plant is 10 meters away from the robot”, “The plant is far away from the robot” or “The plant is by the window”. Such expressions either embody the distance relation between entities by specifying a quantitative measure (“10 meters”) or a qualitative measure that either suggests a long distance (“far away from”) or a short distance (“by”). The corresponding subconcepts of SpatialDistanceModality are therefore QuantitativeDistance and QualitativeDistance; the latter concept then has the two further subcategories Distal and Proximal.

Other, not directly distance-related expressions can imply a certain amount of space between entities too: NonProjection-Axial, also a subclass of Disjointness, captures this distance information. This category specifies expressions that indicate a (relative or absolute) direction according to a specific axis without presupposing a viewpoint-based projective relationship.

¹⁶ The concrete meaning of the expression “a bit further” that describes the extent in the given example is not decomposed further here; it is subsumed under the GUM concept QualitativeSpatialTemporal (see below).

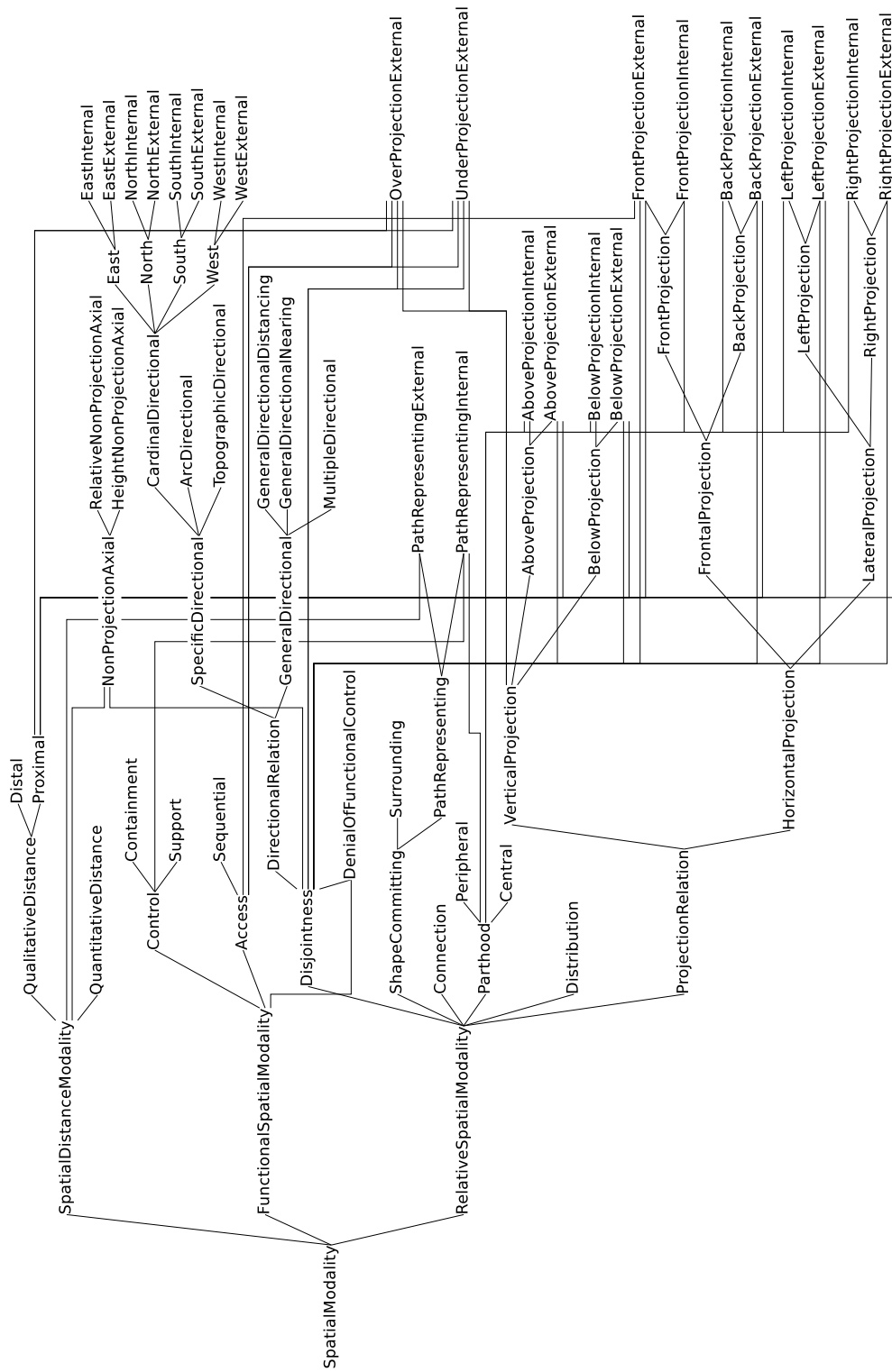


Fig. 7. Spatial modalities in the GUM concept hierarchy.

NonProjectionAxial subsumes the concepts HeightNonProjectionAxial, which refers to expressions construing perpendicular positions, such as “It is deep down”, and RelativeNonProjectionAxial, which refers to relative directions such as “They’re beyond the dining room table”. Moreover, all relationships referring to external projections and external representation of pathways (see below) necessarily imply distance between the locatum/actor and relatum.

4.5.2. Spatial relationships about functional aspects

A FunctionalSpatialModality subsumes expressions about spatial–functional relations between entities. The functions concern accessibility, motion control, support and containment or their opposites, as described by Coventry, Carlson, and others [45,36]. Expressions of this kind of spatial–functional relation mostly go together with distance or relative information between entities. In GUM, FunctionalSpatialModality has the subconcepts Access, Control and DenialOfFunctionalControl.

Access represents relations implying that an entity is physically accessible from another entity, possibly for some ‘purpose’ as intended by this entity (cf. Pustejovsky’s ‘telicity’ field in lexical qualia structure: [133, p. 331]). One of its subclasses is Sequential, which reflects spatial relationships that rely on countable or temporal aspects, as in the example “The library is after the shops” where the relationship between locatum and relatum is reflected by a sequential dependency, i.e., it takes a certain amount of time to move to the placement (“library”). Control represents relations holding between two spatial entities such that one entity controls the position of another entity in space: i.e., if the first entity moves, then the second moves as well. However, the entities do not necessarily have physical contact with each other. Subconcepts of Control are Support and Containment. Support indicates that the control of one entity is caused by an external (though not necessarily physical) contact between the entities, for instance by the presence of gravity (e.g. “The USB drive is on the table”). Containment indicates that the control is caused by (functional) containment (e.g. “The USB drive is in the bag”). DenialOfFunctionalControl represents relations implying that no spatial–functional relation or contact exists between the entities (e.g. “The USB drive is outside the room”, “The lion is out of the cage”).

4.5.3. Spatial relationships about relative aspects of entities

RelativeSpatialModality subsumes expressions about relative spatial relationships between entities. Such expressions imply the involvement of spatial axes (ProjectionRelation), the topological layout between entities (such as Parthood, Connection and Disjointness), restrictions to complex relata (Distribution), or specific kinds of shapes of entities (ShapeCommitting).

ProjectionRelations are divided along the natural three-dimensional spatial axes, i.e., language constructs horizontal (namely lateral and frontal) and vertical directions [168,162]. Projective spatial modalities in GUM are accordingly HorizontalProjection, with subconcepts FrontalProjection and LateralProjection, and VerticalProjection. FrontalProjection is further subdivided into BackProjection and FrontProjection, corresponding to expressions such as “The board is behind/in front of the sofa”, LateralProjection is subdivided into LeftProjection and RightProjection, corresponding to expressions such as “The board is to the left/right of the sofa”. The modality VerticalProjection distinguishes between AboveProjection and BelowProjection, such as “Birds are flying above/below the clouds”, and UnderProjectionExternal and OverProjectionExternal, such as “Birds are flying over/under the clouds”; these concepts are then further differentiated according to their functional commitments [46].

The most specific projective concepts LeftProjection, RightProjection, FrontProjection, BackProjection, AboveProjection, and BelowProjection are also cross-classified according to whether they indicate external or internal relationships between entities with respect to topological positions. For example, FrontProjection can cover both ‘in front of the car’ (external) and ‘in the front of the car’ (internal). External projective modalities are additionally subsumed under Disjointness and Proximal, while internal projective modalities are additionally subsumed under spatial Parthood.

The spatial modality DirectionalRelation covers expressions concerning spatial directions; in many cases these fill source and destination relations. The concept subsumes the two subconcepts GeneralDirectional, subdivided into GeneralDirectionalDistancing and GeneralDirectionalNearing, and SpecificDirectional, representing spatial relations that define directions between entities independently of additional references (e.g. “out of the house”) or depending on additional references, respectively. Such additional references are, for instance, geographic information, as specified by the CardinalDirectional categories and the TopographicDirectional category [110]. Arbitrary directions (e.g., “to and fro”) are specified by MultipleDirectional, while clock directionals (e.g., “The fridge is at three o’clock”) are specified by ArcDirectional. Different languages tend to construct this semantic area rather differently, most traditionally using conventionalized topographical landmarks, such as ‘uphill’/‘downhill’, ‘towards the sea’/‘away from the sea’, etc. [110, p. 148].

As remarked above, some spatial relations require their relata to consist of either a complex object or multiple objects. Linguistically, the relatum then appears either expressed in the plural or as a collective term. These kinds of spatial relations are covered by the modality Distribution. Examples are “It is in the middle of the boxes” (plural) and “He stood among the crowd” (collective). Collective terms are often associated with the GUM element DecomposableObject, i.e., a concept made up of different identifiable parts.

Spatial locations that construe specific parts of their relata are defined by the concepts Peripheral and Central, both subconcepts of spatial Parthood. They construct positions in the middle or at the edge of places, such as “The deer is in the middle of the field” (Central); this example is again an illustration that the distinctions made in GUM are not based on lexical entries but rest instead on grammatical distinctions. The phrase “in the middle of” can be an instantiation of Central as well as Distribution.

Finally, ShapeCommitting refers to expressions in which the relatum is restricted to a specific shape as discussed in Section 3. There are two subconcepts: PathRepresenting and Surrounding. PathRepresenting defines spatial relations that

depend on the shape or the surface of a relatum that expands in at least one dimension; the locatum's position is then described with respect to this shape, or the actor moves along this shape. The concept Surrounding defines spatial relations between entities in which the relata are located somewhere around the locatum or actor. An example of such an expression is "Hedges enclosed the flower beds". Surrounding is also a subconcept of Access.

4.6. Spatial elements III: SimpleThings

The entities that may fill relations referring to participants in a spatial relationship, such as locatum, actor, actee, and relatum, are generally instances of the element SimpleThing. This element represents entities constructed by language as things, and so are generally expressed linguistically by nominal phrases. They can be abstract or concrete, mental or physical, decomposable or non-decomposable, conscious or non-conscious, all of which are subconcepts of SimpleThing. SimpleThings also subsume those elements occurring in spatial modifications, such as QualitativeSpatialTemporal and QuantitativeSpatialTemporal (cf. Section 4.5). These represent qualitative or quantitative spatial (or temporal) information for a spatial modality, such as the qualitative modification "slightly" in "turn slightly to the right", or the quantitative information "45°" in "turn 45° to the right". They are connected with the spatial modality by filling the relation spatialExtent. Moreover, as seen in Section 3, the particular combination of spatial modalities and types of entities can itself bring further constraints on those entities' spatial construal (cf. [7,24]).

Certain other thing-like entities also make specifically spatial commitments that require modeling in a way that appropriately captures their interaction with the spatial modalities they are used with. These are often carried by groups of individual lexical items and so lie at the very edge of what we include within GUM—although there is also usually grammatical evidence to be found from the particular constructions within which they appear. We illustrate three distinct kinds of such classes here.

First, most lexical items for physical objects bring with them commitments concerning the relative extent of their dimensionalities, as described in considerable detail by Lang (e.g., [105]): this is used to motivate distinctive usage patterns such as 'the *height* of the man' vs. 'the *length* of the shadow'. Although not yet present in GUM, an account such as Lang's would clearly be a useful and appropriate extension. Second, there is a further set of lexical items that may be beneficially described as *nominalizations* of spatial modalities, quite analogously to nominalizations of other categories, typically configurations, which are already covered in our generation grammars (cf. Section 5.1). Examples of these kinds of entities include spatial 'parts', such as *interior*, *top*, *bottom*, *edge*, *center*, *border* and so on. Considering these as nominalizations allows them to share the semantic accounts given for their respective spatial modalities and to avoid additional separate treatments within GUM itself. Third, there are also kinds of spatial 'features', such as *bump*, *hole*, *corner* and so on. These are strongly shape committing, although the particular shapes committed to are quite varied as is usual when we move towards the open-class, lexical end of linguistic information.

Both these latter types of spatial terms can be refined further for particular domains of application: for example, *coast* is a specifically geographic kind of edge, and *mountain* might be seen as a specifically geographic kind of 'bump'. Finer classifications of these kinds of entities remain for the future: it will be particularly interesting to compare existing proposals for classificatory features with those predicted from the spatial modalities hierarchy.

4.7. Conclusion

In this section we have presented the organization and main concepts and roles of our spatial extension for the Generalized Upper Model. As we shall see in the evaluation in Section 6, these possibilities cover a substantial proportion of the spatial expressions found in our test corpora, although there is still need for extension and comparison with the situation for other languages (cf. [6,76,77,153,155]). At the present time the GUM ontology as a whole consists of 270 concepts and 110 relations. Of this total, the spatial fragment as described here contributes 98 concepts (36%) and 28 relations (25%), thus forming a significant extension in scope and detail over the original GUM.

The spatial extension has had little effect on the DL expressivity of GUM. The core ontology of GUM is formalized in \mathcal{ALCHN} , i.e., its logic allows role hierarchy (\mathcal{H}), cardinality restrictions (\mathcal{N}), and \mathcal{ALC} -related operators (namely atomic negation, universal restrictions, concept intersection, limited existential quantification, and class complements). The spatial fragment is slightly more expressive than this core part by virtue of introducing data valued roles (also called datatype properties in OWL) and qualified cardinality restrictions: for instance, the spatial extension defines a relation *reciprocalRelation* with a boolean data range, which specifies whether a spatial relation is reciprocal; furthermore, the ranges of some of GUM's spatial roles are not defined by subclasses of Thing, such as the role *orientationRoute*, which is filled with an instance of *GeneralizedRoute* that cannot define a *pathPlacement* or a *pathIndication*. As a consequence, the spatial extension of GUM is expressed in the logic $\mathcal{ALCHQ}(D)$ and remains, as we shall see in the next section, a broad practical linguistic ontology appropriate for a range of natural language processing tasks. The ontology as a whole has been validated as a formally consistent specification; this follows relatively straightforwardly from the fact that GUM contains virtually no cases of negation and so the resulting description is more or less guaranteed to be formally consistent. This has been verified for the current version using Protégé¹⁷ with the Pellet DL reasoner.¹⁸

¹⁷ <http://protege.stanford.edu/>.

¹⁸ <http://pellet.owidl.com>.

5. Application and use of the GUM spatial extension

We have now set out in some detail the categories we propose for a linguistic ontology of space and their interrelationships. The question now arises as to how we can assess and evaluate this proposal's applicability and adequacy. In this section, we address this issue from the perspective of *use* of the ontology. We will consider how the ontology can be beneficial for applications and how it contributes to the construction of semantic interpretations.

5.1. Context of use

In our own work on spatial language, we employ GUM and its spatial extension as the semantic layer in a complete natural language dialogue system enabling interaction both with autonomous robots and with spatial assistance systems providing way-finding or navigation services (e.g., [137,139,138]). For the purposes of the present paper, we can regard the architecture of the system as fairly traditional: speech or written input is passed on to an analysis component that produces semantic representations, these semantic representations are related both to contextual information and to the current state of the dialogue, certain problem-solving tasks are triggered (e.g., finding routes, locating services, etc.), and the results are converted back into a semantic representation, which is then passed back through a generation component to produce either spoken or written output. The problem-solving components of our system represent domain knowledge using ontologies expressed in description logic and perform their tasks using the well-known description logic reasoners Pellet¹⁹ and Racer,²⁰ some application-specific implementations of navigation and route-finding algorithms, as well as further dedicated spatial reasoners.

The role of GUM in this environment is to support natural language interaction by defining the semantic types and constraints that hold at the semantic level of analysis: instantial information of the type used for illustration in the previous section is produced by the analysis component and triggers contextualization. However, in contrast to many systems where contextualization is seen as a process of further specification filling in the underspecified semantics derived from the input [91,88], with GUM we allow for a much broader range of 'inter-ontology' alignments between the linguistic ontology (and instantial information with respect to that ontology) and the ontologies of the application (and instantial data with respect to these). It is in this latter approach to contextualization that much of the flexibility of natural spatial language usage is achieved.

Our automatic generation component, taking semantic specifications formulated in terms of GUM to surface strings, follows the long established techniques for using the Penman Upper Model in natural language generation (cf. Section 2.3 above), relying on broad coverage systemic-functional generation grammars for English and German [17]. More interesting for the current paper, however, is the approach we take to analysis, since (a) an Upper Model-based semantics has not previously been coupled with an analysis component and (b) the requirements of analysis demand a more careful consideration of semantic compositionality during interpretation. The basis of our analysis component also turns out, nevertheless, to be relatively straightforward. We have adopted the framework of combinatory categorial grammar, and more specifically the version producing Hybrid Logic Dependency Semantics (HLDS) developed by Kruijff and Baldridge [11]. HLDS specifications are sorted hybrid modal logic representations, in which the sorts are provided by semantic type hierarchies defined along with the grammar. The grammars themselves are highly lexicalized and the semantics of lexical items use the same defined semantic types. Semantic interpretations of utterances are built up by functional composition in the usual way. For performing the analysis itself we employ the OpenCCG toolset.²¹

Our initial analysis grammars were constructed exactly parallel to the spatial language analysis grammars developed by Kruijff and colleagues [96] with the single exception that the semantic types employed in lexical entries are, in our case, drawn from the GUM ontology. The HLDS analysis results of Kruijff et al., in contrast, employ semantic types adopted from a combined linguistic-domain commonsense ontology. This is a practical solution in their case because the domain of application is relatively fixed and the spatial language that occurs accordingly limited in scope. It would, however, be interesting to consider a GUM-based semantic layer in the context of their system also.

The final connection to our discussion here lies in bridging between HLDS specifications and ABox assertions within description logic. Due to the close formal relationship existing between modal logic and description logic [11], it is in fact straightforward to consider the output of our grammars produced using OpenCCG as ABox assertions of exactly the kind we have given in our examples above. We show this very briefly here as follows. Our grammars contain lexical items of the form depicted in Fig. 8 for spatial prepositions such as *in*, *on*, etc. On the left-hand side of the figure we find a complex syntactic category and on the right-hand side its corresponding semantics. This particular lexical item then states that one syntactic reading of 'in' is as a category that is 'looking' for an NP on its right-hand side (indicated by the forward slash) in order to construct the complex category NP\NP—i.e., a category that is itself looking for an NP on its *left* (indicated by the backslash) to yield an NP: in other words, an NP post-modifier such as a PP.²² The semantics for this category includes the

¹⁹ <http://pellet.owldl.com>.

²⁰ <http://www.sts.tu-harburg.de/~r.f.moeller/racer>.

²¹ <http://openccg.sourceforge.net/>.

²² We should note that our grammars are still under development and so nothing rests upon the particular categories given here: the mechanisms for composition are not affected.

$$\frac{in}{np \setminus, np /, np} : \begin{array}{l} @x1:gs-GeneralizedLocation \\ \quad (<gs-hasSpatialModality>(x2:gs-Containment \wedge in) \\ \quad <gs-relatum>x4:gum-SimpleThing) \wedge \\ @x6:gs-SpatialLocating \\ \quad (<gs-locatum>x5 \wedge <gs-placement>x1) \end{array}$$

Fig. 8. One of our OpenCCG lexical entries for the spatial preposition *in* as an NP modifier. The prefix *gs-* stands for ‘GUM space’, symbols before colons are variable names, the at symbol (@) marks HLDS named ‘nominals’, the caret (^) is logical conjunction, and the symbols within angle brackets correspond to roles.

<pre>(@x1:gs-SpatialLocating (<gs-locatum>b1:slm-Box ^ <gs-placement> (x2:gs-GeneralizedLocation ^ <gs-hasSpatialModality> (o1:gs-Support ^ on) ^ <gs-relatum>s1:slm-Shelf)) ^ @x3:gs-SpatialLocating (<gs-locatum>s1:slm-Shelf ^ <gs-placement> (x4:gs-GeneralizedLocation ^ <gs-hasSpatialModality> (i1:gs-Containment ^ in) ^ <gs-relatum> k1:slm-Kitchen)) ^ @b1:slm-Box)</pre>	<pre>(@x1:gs-SpatialLocating (<gs-locatum>b1:slm-Box ^ <gs-placement> (x2:gs-GeneralizedLocation ^ <gs-hasSpatialModality> (i1:gs-Containment ^ in) ^ <gs-relatum>k1:slm-Kitchen)) ^ @x3:gs-SpatialLocating (<gs-locatum>b1:slm-Box ^ <gs-placement> (x4:gs-GeneralizedLocation ^ <gs-hasSpatialModality> (o1:gs-Support ^ on) ^ <gs-relatum>s1:slm-Shelf)) ^ @b1:slm-Box)</pre>
--	--

Fig. 9. The HLDS specifications produced with our grammars for the phrases in ((26)a: left-hand side) and in ((26)b: right-hand side). The prefix *slm* stands for ‘semantic lower model’, a semantic lexicon defined for the domain.

two semantic types necessary for this reading: first, a *GeneralizedLocation* corresponding to the NP\NP (or PP), providing a semantic ‘place’; and second, a *SpatialLocating* that has the place as its placement and an empty slot, *x5*, for the object to be located. This is then directly comparable with the single-place predicate interpretation given in Section 3.1 above. The *GeneralizedLocation* contains a spatial modality appropriate for ‘in’, i.e., *Containment*, which is classified within GUM as one of two kinds of functional control (cf. Fig. 7). The full form of the entry also links the semantics of the middle NP with the variable *x5* (the *locatum*), since this is the NP being sought on the left, and the semantics of the rightmost NP with the variable *x4* (the *relatum*); in short, the semantics corresponds to that of a phrase [*x5* in *x4*]_{NP} just as described in Section 3 above.

The work of constructing complete semantic specifications is then carried out by functional application, modelled in OpenCCG by unification. Fig. 9 shows the semantics that result when the phrases in (26) are parsed with our current grammars.

- (26) a. the box on the shelf in the kitchen
 b. the box in the kitchen on the shelf

Naturally, many more parses are produced but those shown are the first two returned for each phrase; we also delete here all attribute values that are not concerned with spatial information. The most significant difference between the two is that quite different spatial relationships have been constructed. In the first specification, it is the shelf that is in the kitchen (*SpatialLocating* *x3*), whereas in the second it is the box that is in the kitchen (*SpatialLocating* *x1*). We have tuned the grammars so that more likely interpretations are produced first: in the present case, it is less likely that one finds the kitchen on the shelf—although, of course, this is still semantically possible (the kitchen may be a doll’s house kitchen for example) and so is still present as a less favored parse. The correspondences with the Manchester syntax specifications used elsewhere in this paper should be self evident.²³ There is still substantial work to be done in selecting the most contextually appropriate parses from the solution set returned; however, our concern at present is simply to restrict that set to spatially possible parses.

For more straightforward utterances our grammars are already showing useful results. Although these grammars remain experimental, we have benchmarked them against two testbeds of spatial language constructions. The first testbed was built

²³ This is a straightforward syntactic translation. Within our dialogue system we use the XML form of output from OpenCCG and convert this via XSLT to a syntax appropriate for the DL reasoners we use.

directly from the set of spatial language constructions selected to illustrate each concept within the complete GUM spatial extension documentation [83], while the second testbed is derived from the spatial language corpora used in the evaluation of linguistic coverage that we describe in Section 6.1. At the time of submission the English grammar provides parses for 98% of the first testbed, and just over 50% for the larger corpus-derived testbed. As with all our resources, the analysis grammars are freely available for use and extension.²⁴

Although hand-crafted grammars of this kind have substantially smaller coverage than wide-coverage statistical parsers, we believe that their accuracy and quality is significantly higher with respect to spatial language than existing statistical parsers. Statistical approaches neither provide a detailed account of spatial meaning, nor take spatial semantic constraints into account in the parsing process. That notwithstanding, the application of the GUM ontology to existing wide coverage linguistic resources also poses a challenging and useful direction for future research. Moreover, while the manual annotation of spatial data with respect to proposed ontological categories described in Section 6.1 is a useful measure of overall model coverage, it is evident that a far better indication of coverage could be obtained if analysis were automated.

5.2. Representational and semantic levels: Adequate inferences

The provision of a linguistic ontology makes possible the performance of a particular range of inferences, both internally to the ontology and with respect to the modules and other layers of information with which it is intended to interact. Another argument for our decomposition of the spatial semantics problem is that different kinds of inferential work can be distributed more appropriately. In this subsection, we set out this role of our linguistic ontology in more detail.

5.2.1. Compositionality and interpretation

We have seen in this section how instantial semantic representations of utterances (represented as ABox assertions within description logic or, equivalently, as HLDS specifications) can be constructed employing an appropriate computational analysis grammar. We consider in this subsection where this stands in the entire chain of interpretation that we require for a situated spatial dialogue system and spatial contextualization. We can consider the ability to construct appropriately structured semantic specifications anchored against the semantic terms of the linguistic ontology as a rather weak kind of compositionality. The semantic specifications of complete phrases are composed from the semantic specifications of their components, respecting the constraints on concept construction specified in the ontology. But for further contextualization, it is necessary to consider what these semantic specifications mean in their own right.

We stressed in the opening sections of the paper how we were committed to an ‘indirect’ approach to spatial semantics, one which resists immediate interpretations of linguistic spatial expressions in terms of, for example, regions of three-dimensional space, in order to avoid over-commitment and inflexibility in interpretation. We therefore see GUM semantic specifications as performing the following central task in building interpretations. Information derived from utterances is organized according to GUM into specific relationships and categories. The function of this organization is to provide anchor points to which constraints on appropriate kinds of (non-linguistic) spatial semantics can be attached. These constraints are seen as ‘minimal’ properties to be met by (non-linguistic) spatial formalizations of particular linguistic semantic categories rather than specific accounts of space ‘as a whole’. Flexibility in interpretation then results from the fact that a variety of different spatial theories might be adopted, as long as the minimal conditions on what is to be expressed for a particular category are met.

We will illustrate this approach by example, showing how it opens up several important directions for future research. In particular, we can usefully relate the framework to earlier discussions of the so-called ‘two-level approach’ to linguistic semantics (e.g., [164,165,104]), in which a compositional semantic level is separated from a non-compositional ‘conceptual’ level. We also seek to maintain a compositional semantic level but go further by applying newly emerging results in ontology structuring and inter-ontology alignment to explore the relationships between levels within a more formal setting [100–102, 99]. Detailed discussion of this component of our work goes beyond the scope of the present paper (see [81,82,80]), but our examples will suggest the line of development we are following. It should, however, be noted here that it is only with an ontology-like specification of the linguistic semantics, such as that offered by GUM for example, that the tools of inter-ontology alignment become applicable to this problem at all. We also see this as an important contribution of our approach.

Most of the more formal approaches to spatial language introduced in Section 2 above have also aimed for compositionality. We saw there that some choices of spatial formalization are less suitable than others. The geometric (time-dependent) three-dimensional space interpretations of Kracht and of Francez and Steedman [94,58], for example, restricted the scope of their accounts to what we have termed ‘geometrically transparent’ usages. Extending beyond this, the treatment of modification in spatial prepositional phrases proposed by Zwarts [168]—such as “*one meter* behind the desk”, “*right* under the lamp”, “*far* outside the village”—provides further convincing evidence that a compositional approach based on three-dimensional regions cannot be sufficient since information necessary for the construction of interpretations of these modified phrases is then simply not available. In particular, to support a composition in which the semantics of *one meter* and the semantics of *behind the desk*, or the semantics of *right* and *under the lamp*, can be combined, it is necessary for the denotations of

²⁴ <http://www.diaspace.org/grammars>.

the second members of each pair to include *distance* and *direction* information. This is not immediately available from a set of points making up a region. Zwarts argues that the appropriate denotation for spatial expressions is therefore to be found in sets of vectors pre-anchored in the specified reference objects. Distance modification then constrains the length of the vectors considered, and direction modification constrains which out of a broader range of vectors at issue is relevant according to angle.

This perspective is useful because it emphasizes an important precondition for compositionality as such: if the information necessary for a composition is not available in the entities being composed, then it will not be possible to construct a compositional account: whatever information is placed within the semantics of our linguistic spatial entities, we need to ensure that it is sufficient for the meanings that are to be constructed with it. We can expect, therefore, that further properties beyond the sets of vectors that Zwarts proposed will also be required: for example, functional notions may support modifications of the kind “*properly* on the table”, “*right* over the brush” and so on: this suggests that telic notions will also belong in the semantic specification as proposed for lexical semantics in general by Pustejovsky [132]. Patterns of modification are then also shown to offer grammatical syndromes relevant in the construction of our semantics.

Different areas of the linguistic ontology appear to require support from different kinds of abstract spatial theories. In Section 2.2 above, we briefly mentioned Eschenbach’s account of ‘front’/‘behind’ and ‘left’/‘right’ in terms of half-lines and half-planes respectively [53,54]. To the extent that this characterization is further supported by empirical studies, we associate such distinct abstract theories with GUM categories in order to provide a summary of the abstract spatial commitments of those categories: for example, the spatial modality *LateralProjection* would then be connected with a theory that at least defines half-planes. The subconcepts of this category within GUM, i.e., *RightProjection* and *LeftProjection*, are then associated with theories formally extending that of the superconcept. Precisely which kinds of extensions are called for in each case, for example whether they are conservative (cf. [4,101]) or not, itself defines several important further research questions. A similar situation holds for the spatial modality *FrontalProjection* and its subcategories *BackProjection* and *FrontProjection*—this area would, given appropriate support, be related to a theory of half-lines in the same way.

The entire area of routes, destinations, sources, etc. described in the previous section under the concept *GeneralizedRoute* requires a different treatment. The distinctions drawn here on grammatical evidence appear to be covered well by Zwarts’ proposals for an algebra of paths [169], related to his notion of sets of vectors mentioned above. Zwarts’ use of linguistic aspect in his argumentation also illustrates another powerful grammatical syndrome: the inter-relationship of notions of boundedness and temporality. In addition, certain properties of paths that he specifies algebraically provide important guidelines for interpreting the distinct combinations of routes and paths specified in GUM. For example, we distinguished in Section 4 above between location expressions consisting of components that co-describe a route—for example, by giving its source and destination—and expressions that describe a sequence of path segments. The GUM specification of these is quite distinct and allows us to associate their interpretations with path intersection or path concatenation respectively. Zwarts’ results concerning constraints on aspect are then directly importable, as we will illustrate in the subsection following.

The bundling of source, destination and path information within the GUM concept *GeneralizedRoute* also makes contact with the body of work drawn on by Kracht [94]. His formal distinction between modes (concerning movement) and localizers (static relationships) corresponds loosely with GUM’s division between roles of the *GeneralizedRoute* (Kracht’s modes) and the places filling those roles (Kracht’s location phrases). Similar constructs are given by Asher and Sablayrolles in their account of French motion verbs [6]. Moreover, we see here clearly the value of maintaining a looser coupling between GUM specifications and corresponding spatial theories since the approaches to semantic interpretation taken by Kracht and by Asher and Sablayrolles are significantly different. For Kracht, we have a normal case of compositionality by which the meaning of the mode (the directionality) is composed with that of the location phrase; for Asher and Sablayrolles, we have instead many cases of nonmonotonic discourse interpretation. Both approaches have beneficial properties and it is by no means clear at this time whether one account is always to be preferred to the other. However, *regardless of the account we select*, the same kinds of distinctions will need to be drawn on the basis of the linguistic utterances considered. These distinctions may then be interpreted further in various ways, but the grammatical syndromes drawn upon for their formulation will not change. Thus we can strongly presume that any account proposed will need the distinctions drawn by GUM; moreover, any account drawing on these distinctions receives as a by-product a well specified mapping down to surface form.

To show these various features of our account in slightly more detail, we turn from an abstract comparison to one concrete case of compositional semantic interpretation drawing on Kracht’s description. This illustrates our suggested use of GUM as a collection point for semantic constraints. As already mentioned, within Kracht’s account the semantics of modes combines with that of location phrases. The semantics of localizers, i.e., spatial prepositions, are taken to be functions that take a referent object (our *relatum*) and a time, and which produce the set of regions standing in some specified geometric relationship with the reference object at that time [94, p. 188]. The semantics of modes is then constructed essentially by decomposing temporal intervals in order to characterize when propositions hold. For example, the notion of a ‘coinitial’ mode insists that some spatial configuration obtains at the ‘beginning’ of the period over which an event is described. Thus, in Kracht’s sentence:

(27) The cat came out from under the table.

we have a (static) location phrase “under the table” that Kracht takes as denoting the set of three-dimensional regions geometrically ‘under’ the table at any given time and the mode for ‘from’ picks out temporal intervals in which the location

phrase is true of the mover. The result is then just those events where the selected intervals strictly begin the interval of the event as a whole: i.e., the cat started out under the table and ends up not under the table. The multiword lexeme ‘came out’ then provides a set of events with which this semantics must be combined. As with Francez and Steedman’s account, the semantics of the spatial terms here incorporates many items that are not strictly to do with spatial language but which serve to bind together information during composition; these are also specific to the fine-grained detail of the account as a whole.²⁵

A description of the same sentence drawing on GUM’s categories (and which would be produced by automatic analysis) can be seen by analogy to example (24) above; this description involves a *NonAffectingDirectedMotion* with an actor of a cat and a route. The filler of the route role is supplied as usual by the spatial prepositional phrase, in this case: “from under the table”. The ‘from’ is picked up as a specification of the source of the route, also as usual. Now, when we turn to the embedded phrase “under the table”, we have a difference to our earlier example. The spatial modality of a *GeneralizedLocation* filling the slot of a source depends on the syntactic category of the linguistic unit expressing that *GeneralizedLocation*. When this is a noun phrase (e.g., “from *the hill*”), the spatial modality is provided by the source as a *GeneralDirectionalDistancing*; this is analogous to a coercion or the interpretation of an empty place-indicating ‘preposition’ in those accounts that strongly bind syntactic structure to semantic structure (cf. [95]). When the embedded unit is a prepositional phrase, however, it contributes its own spatial modality directly. The GUM specification of the embedded PP here is also constructed as normal, yielding for “under the table” a *GeneralizedLocation* with *relatum* “table” and a *hasSpatialModality* of *UnderProjectionExternal*, a subcategory of *VerticalProjection*, *Disjointness* and functional *Access* (cf. Fig. 7). Each of these superconcepts contributes spatial theories for further interpretation; the vertical projection may be appropriate for a half-plane description in the style proposed for ‘left’/‘right’ by Eschenbach above (with the addition of gravity as a potential orientational anchor) or as a set of (downward pointing) vectors as proposed by Zwarts, disjointness may be covered by a spatial mereology or region calculus, and access by some notion of ‘protection from’ (as in “under the umbrella”).

For the semantic interpretation of the combined phrase, we combine the contributing semantics of source and its *GeneralizedLocation* accordingly. Our general ontological approach is committed to high degrees of modularity, leading us to maintain domains as separate as far as possible. This is most compatible with the treatment of Zwarts, who also attempts to keep spatial semantics purely spatial. With an association of *GeneralizedRoute* with Zwarts’ path algebra, we can take the denotation of the current route as the set of paths originating in the place given by the source. But, alternatively, with an association with Asher and Sablayrolles’ spatial regions relevant for movement, we could similarly have a path from a specified starting area. The GUM backbone specification tells us *which* semantic contributions are to be combined, but does not demand conformance to one particular non-linguistic spatial construal rather than another.

The final part of the semantic interpretation is carried by the additional fact that the *GeneralizedRoute* fills the role of a route in the *NonAffectingDirectedMotion* spatial configuration as a whole. This calls for the information to be combined and, again, several options are available. Possibilities include taking the intersection of events from Kracht, the relation of two domains (the spatial ‘trace’ of the event and the path itself) from Zwarts, or the simpler logical conjunction of Asher and Sablayrolles axioms for verbs of movement.

In a sense, the specifications produced with respect to the linguistic ontology provide a skeleton for compositional semantic interpretation that generalizes both across differing syntactic treatments and across different theories of space. We do not believe that it is possible at this time (and perhaps it will never be possible) to decide once and for all on *the* spatial theory appropriate for all natural language spatial semantics in all contexts of use. As a consequence, we consider the provision of a linguistic ontology such as GUM as an essential mediator between approaches and between levels. Different contexts of use may well demand spatial theories with differing formal properties; with the associations to abstract spatial theories pursued with respect to GUM, we therefore seek to provide explicit specifications of the minimal requirements that need to be met by a spatial theory when particular kinds of spatial language are to be used. In subsequent phases of interpretation, we then relate statements made in these spatial theories to non-linguistic, domain-oriented descriptions of space, which may also be ontological in nature. This approach also avoids the danger of needing to bring ever more non-linguistically motivated information into the shallow semantics in order to provide for later compositionality.

5.2.2. Internal inferential relationships within GUM

In the previous subsection we saw the role of GUM as offering a guide for subsequent semantic interpretation. In this subsection we consider the inferential possibilities provided by GUM itself. Most obviously, the GUM ontology can be used to analyze whether a specification of a particular sentence satisfies certain conditions and requirements. For instance, a direction can only be defined by a *NonAffectingSpatialDoing* or *AffectingSpatialAction*, but not by a *SpatialLocating*. Such dependencies and constraints can be automatically verified in terms of ontological consistency of the sentence instantiation (ABox consistency) using standard description logic reasoners. They then also feed into syntactic analysis via semantic type checking during composition. This information is important for subsequent interpretation because it makes it possible to identify semantic gaps and constraints that raise explicit goals for pragmatic contextualization.

We exemplify four types of linguistic inference supported by GUM specifications: two for utterances taken from our corpus, an invented one inspired by arguments put forward by Kracht [94], and a final set of examples showing how

²⁵ For some further discussion focusing on linguistic aspects of the accounts of Kracht, Francez and Steedman, Asher and Sablayrolles, and Zwarts in comparison to our own, see Bateman [13].

the inherently limited semantics possible within GUM also has the positive effect of blocking inference chains that more powerful formalizations might wrongly suggest as plausible.

We start with example (28) in order to show both what information is entailed and where pragmatic inferences must take this further.

(28) neunzig Grad Rechtsdrehung
ninety degrees rotation to the right

The corresponding GUM specification is as follows (the specification of *angleDegrees*, an instance of QualitativeSpatial-Temporal, is omitted for simplicity):

<p>(29) Individual: NAOC Types: NonAffectingOrientationChange Facts: actor <i>undefined</i> processInConfiguration <i>rotating</i> orientationDirection <i>GenLoc</i></p>	<p>Individual: GenLoc Types: GeneralizedLocation Facts: relatum <i>undefined</i> hasSpatialModality <i>RightProjection</i> quantitativeAngleExtent <i>angleDegrees</i></p>
--	---

Although semantically unambiguous, this phrase is underspecified with respect to its contextual interpretation. Nevertheless, the phrase succeeds in evoking a range of realistic expectations about what kind of spatial movement is being described. The GUM specification captures semantic entailments that drive these intuitions as follows.

First, the configuration as a whole is defined as NonAffectingOrientationChange. This implies (going down the hierarchy in GUM to the corresponding branch) that:

- (i) DoingAndHappening: Somebody is doing something or something is happening. This event involves at least a processInConfiguration, which is here defined as 'drehen' (rotate).
- (ii) NonAffectingAction: The event doesn't affect anybody or anything; there is no actee in this configuration, whose role would otherwise need to be filled.
- (iii) NonAffectingDoing: The event involves an actor who is doing something. This role is not given explicitly in the configuration and can be formally identified as missing.
- (iv) NonAffectingSpatialDoing: The event is a spatial action; therefore it may define a placement where the action takes place: this is not the case here and so we know that this information although potentially relevant is not being provided directly.
- (v) NonAffectingMotion: The actor performs a spatial movement, namely one of three possibilities: directed motion (route or direction), orientation change (orientationDirection or orientationRoute), or simple motion (placement).
- (vi) NonAffectingOrientationChange: The movement performed by the actor is a change in orientation, defined by the relation orientationDirection.

The relation orientationDirection is filled by a GeneralizedLocation, which requires at least a spatial modality and a relatum. While the relation hasSpatialModality is defined as RightProjection, the slot for the relation relatum remains unfilled ('turn to the right of what?').

By this point, therefore, we have identified where specific information needs to be added by pragmatic inference for full contextualization. For example, one default relatum for the right direction in such a situation would be the actor, i.e., the actor would turn to their own right. Further pragmatic information may be derived from the (linguistic or situational) discourse context; if the actor has been the addressee in previous utterances, it can be inferred pragmatically that this is also true for the current (semantically undefined) actor, who is then animate (derived from the fact of being an addressee), with a high likelihood of possessing intrinsic directions, thus instantiating a likely relatum. Thus, GUM identifies the expected roles and concepts for a particular configuration. At specific points in the hierarchy, the requirements of the GUM concept definitions highlight that both the actor and the relatum are missing. To fully understand the description, the information about who is doing the turning action, and what the direction 'right' relates to, need to be ascertained. The fact that the '90 degree' information is optional suggests that there is some redundancy involved here—a reorientation to the right, for example, can be interpreted even without the quantitative specification.

Our next example illustrates how semantic ambiguity leads to different kinds of possible inferences, again using an example from our corpus.

(30) Go down three inches.

Corresponding GUM specifications are given in (31); we omit the internal definition of the instance *distExtent* as this is not of central concern. We see here that there is already semantic ambiguity to be resolved. The spatial expression 'down' is lexically compatible with at least two distinct readings: one as involving a SpecificDirectional spatial modality, another as a

PathRepresentingInternal. The entailments following from these readings within GUM also differ in ways significant for the interpretation.

- | | |
|---|---|
| <p>(31) Individual: <i>NADM1</i>
 Types: NonAffectingDirectedMotion
 Facts: actor <i>undefined.1</i>
 processInConfiguration <i>going</i>
 direction <i>GenLoc</i></p> | <p>Individual: <i>GenLoc</i>
 Types: <i>GeneralizedLocation</i>
 Facts: relatum <i>undefined.1</i>
 hasSpatialModality <i>SpecificDirectional</i>
 quantitativeDistanceExtent <i>distExtent</i></p> |
| <p>– OR –</p> | |
| <p>Individual: <i>NADM2</i>
 Types: NonAffectingDirectedMotion
 Facts: actor <i>undefined</i>
 processInConfiguration <i>going</i>
 route <i>GenRoute</i></p> | <p>Individual: <i>GenRoute</i>
 Types: <i>GeneralizedRoute</i>
 Facts: pathPlacement <i>GPL</i>
 Individual: <i>GPL</i>
 Types: <i>GeneralizedPathLocation</i>
 hasSpatialModality <i>PathRepresentingInternal</i>
 relatum <i>undefined</i>
 quantitativeDistanceExtent <i>distExtent</i></p> |

Construction of the semantics proceeds as follows. Spatial modalities contribute to GeneralizedLocations (cf. Fig. 5 or Definition (10)). In the first reading, the SpecificDirectional modality suggested lexically allows a GeneralizedLocation to be built directly; the spatial extent of “three inches” is combined as a quantitativeDistantExtent (cf. Section 4.5). In the second reading, the alternative PathRepresentingInternal modality supports a finer classification as part of a GeneralizedPathLocation, i.e., as a segment of a path; the spatial extent information is treated as in the first reading. In both cases no relatum is explicitly supplied by the linguistic form, although semantically required (and hence entailed).

Developing the first alternative further, a GeneralizedLocation may fill a variety of roles, including a placement role in SpatialLocating configurations (cf. (15)), a source or destination in generalized routes (23), a direction role in a NonAffectingDirectedMotion (18), an orientationDirection in a NonAffectingOrientationChange (20), and so on.²⁶ The verb contributes particular lexical information that supports a NonAffectingDirectedMotion analysis and this binds the GeneralizedLocation into the spatial configuration as a direction. Here, although the actor is not specified in the linguistic form, the imperative allows this to be filled with the addressee. Moreover, the relatum of directions is always the actor/mover and therefore is also identifiable by co-reference (indicated in the specification by the extension .1) without further pragmatic interpretation.

The second alternative is more constrained from the outset; a GeneralizedPathLocation can only form the pathPlacement or the pathIndication in a GeneralizedRoute (23). Here, the internal path is not construed as a point and so pathPlacement is the appropriate choice. The lexical semantics of the verb as a NonAffectingDirectedMotion also makes available the possibility of binding a GeneralizedRoute into the configuration with the route role (18) and so this is a further possible semantic analysis for the example sentence as a whole. In this case, the fact that a relatum is left unspecified in the surface form for the spatial modality has consequences differing to those for the first interpretation. Since the required modality is not a direction but a PathRepresentingInternal, this means that the action takes place in relation to an entity that is capable of representing the path; such entities could be hallways, streets, roads, rivers, borders, etc. This information does then need to be filled in pragmatically. The explicit mention of a relatum can disambiguate the two possible interpretations of ‘down’.

In our third example, we illustrate how GUM leads to limited inferences with respect to the ambiguous and pragmatically highly underspecified (though semantically complete) example shown in (32).

- (32) Jack ran in front of the bus.

The corresponding GUM specifications are:

- | | |
|---|---|
| <p>(33) Individual: <i>NAM</i>
 Types: NonAffectingMotion
 Facts: actor <i>Jack</i>
 processInConfiguration <i>running</i>
 placement <i>GenLoc1</i></p> | <p>Individual: <i>GenLoc1</i>
 Types: <i>GeneralizedLocation</i>
 Facts: relatum <i>Bus</i>
 hasSpatialModality <i>FrontProjectionExternal</i></p> |
| <p>– OR –</p> | |

²⁶ In current work we are exploring further the extent to which finer-grained role filler constraints can reduce the potential range of interpretations earlier; this is primarily a combined empirical-semantic analysis task.

Individual: *NADM***Types:** *NonAffectingDirectedMotion***Facts:** actor *Jack*processInConfiguration *running*route *GenRoute***Individual:** *GenRoute***Types:** *GeneralizedRoute***Facts:** destination *GenLoc2***Individual:** *GenLoc2***Types:** *GeneralizedLocation*relatum *Bus*hasSpatialModality *FrontProjectionExternal*

Following through the detailed interpretation in the manner shown in the previous example, this utterance can be interpreted in two basic ways. In the case of a *NonAffectingMotion* interpretation, a spatial action is specified without a direction or orientation; however, the location where the motion takes place is specified via a placement relation: the action of running takes place in front of the bus. In the case of a *NonAffectingDirectedMotion* interpretation, the destination of a route is specified as being in front of the bus. In both cases, this exhausts the possible specifications based on the given semantics of the utterance. A number of inferences can be drawn on this basis. If the utterance is interpreted as being part of a route, this entails that other route elements can be added, namely a source, pathIndications, and pathPlacements.

Consider the following example in which these optional slots are filled:

(34) Jack ran out of the house, through the gate, across the field, and in front of the bus.

Here the alternative reading of a motion event taking place in front of the bus is very much less likely, since the presence of a source ('out of the house') induces a *GeneralizedRoute* interpretation. Furthermore, the fact that a destination constitutes the end point of a motion configuration entails that the action is bounded, while no such inference holds for the placement interpretation. Boundedness is traditionally tested by considering the acceptability and logical consequences of adding various temporal adjuncts [169], as in:

(35) Jack ran in front of the bus for three hours.

Here, only the placement interpretation is available. Adding 'for three hours' to the modified example (34) renders the sentence semantically unacceptable in the intended sense:

(36) * Jack ran out of the house, through the gate, across the field, and in front of the bus for three hours.

In contrast, adding another punctual event is unproblematic for the destination version (example (37)) but unacceptable for the placement version (example (38)):

(37) Jack ran out of the house, through the gate, across the field, and in front of the bus. He was killed immediately.

(38) * Jack ran in front of the bus for three hours. He was killed immediately.

In addition to these semantically based inferences, pragmatic considerations open up a further range of interpretations, all of which are nevertheless perfectly possible given the linguistic material considered. The spatial modality *FrontProjectionExternal* does not indicate, for example, just how the front relation is being projected. For instance, various different perspectives could be used in this situation. The front direction could be defined based on the intrinsic front side of the bus, or on the current direction of movement of the bus (which could be going backwards), or on an onlooker's perspective; each of these yield different regions for 'front' (at least potentially). Furthermore, the given semantics does not have anything to say about whether or not the bus is currently in motion. In the former interpretation, it is additionally unclear whether the process of running is directed or the actor is running in circles; therefore, the specification in GUM remains on the branch above the differentiation of *NonAffectingSimpleMotion* and *NonAffectingDirectedMotion*. All of these details would have to be derived from the discourse context, which is likely to either provide enough information to rule out some of the available alternatives or else render them irrelevant. In short, beyond the inferences described above that can actually be made based on the semantics of the two basic interpretations, there is neither a necessity nor a semantic basis for spelling out the broad range of possible geometric configurations potentially represented by this utterance.

In our last examples, we show how the definitions in GUM can also be used to block invalid reasoning chains that might otherwise be invoked. For example, GUM does not specify projective relations to be transitive. It might be thought that it is a matter of the lexical semantics that one can infer from 'A' being to the left of 'B' and 'B' being to the left of 'C', that 'A' is also to the left of 'C'. But this is misleading. It can readily turn out that, due to differences in granularity or objects arranged in circles, etc. (cf. [81]), that such relations do *not* hold. We need to be very cautious about specifying such information as additional axioms in the linguistic semantics therefore. A more appropriate place to state such regularities is in the abstract representations of space, since here it is possible, for example, to specify precisely transitive and inverse relations among projective relations and to determine when they hold—regardless of linguistic descriptions and their inferences. Many such specifications have been developed within the field of qualitative spatial representation and reasoning [42]—the double-

cross calculus [62], for instance, precisely models orientation relations and their composition. Spatial inferences can then be performed on the basis of the calculus, i.e., it can be analyzed whether ‘left’ is the inverse of ‘right’ *in a given situation*. This has been illustrated with respect to several parts of GUM—for example, the horizontal projective relations are related to the double-cross calculus in Hois and Kutz [81].²⁷ Further informal examples of the link between linguistic semantics and formal spatial theories have been presented by Bateman, Tenbrink and Farrar [19], while examples of the modular formal spatial theories we employ independently of language can be seen in Bateman et al. [14].

There are also interesting interactions with compositionality specifically concerned with what interpretative work is done where. For example, in the approach of Asher and Sablayrolles [6], certain information is pushed down into the lexical items of individual verb families that from our perspective may be better located at the configuration level. Asher and Sablayrolles note, for example, that for a verb such as ‘run’ (again their examples are actually in French but we will only use translations for current purposes) it is not yet possible to fully specify whether movement from a location to another location is taking place:

(39) The man ran to the park.

(40) The man ran on the spot.

To cope with this, they suggest adding a defeasible inference to the lexical semantics of ‘run’ concerning whether or not the eventuality described involves change of location. The corresponding GUM specifications instead move this information up to the configuration: example (39) involves a `NonAffectingMotion` and a destination relation within a route, while (40) is a `NonAffectingDirectedMotion` with an associated placement relation. All that ‘run’ as a kind of Process itself contributes is a manner of moving; the question of whether a location-change is involved is not linguistically constructed outside of the context of an entire clause. Again this means that we can avoid importing defeasible rules into our linguistic semantics.

5.3. Conclusions and discussion

In this section, we have described the use of GUM as part of larger natural language processing systems and the kinds of contributions that it makes to compositional semantic interpretation and subsequent inference processes. We should also compare GUM with other ontologies that contain descriptions of space, particularly those that build on principles of ontological engineering. Ontology already plays an important role here and we find several areas where explicit ontological formalizations have been pursued (cf. [59,38,60,57,52]). Normative standards, such as OGC’s GML (ISO 19136) and OpenGIS, also relate well to ontological accounts but have not been shaped significantly by linguistic concerns. As a consequence they only address half of the language–space relationship, leading to the problems set out in detail in Section 2. Similarly, it is difficult to relate spatial ontologies that define categories of linguistic spatial relations according to ‘conceptual’ schemes (e.g., [143,121]) to GUM without compromising the flexibility of usage characteristic of natural spatial language.

Some general purpose ontologies also contain significant information about space; for example both `ResearchCyc`²⁸ and the `Suggested Upper Merged Ontology` (SUMO: [125]) formalize a variety of spatial theories; Bateman and Farrar offer an extensive overview of this area [15]. The base ontology of SUMO, for example, defines several general spatial concepts. One of them with a substantial number of subconcepts is the category `SpatialRelation` representing many spatial relations based on mereological and topological relations. Although this categorization of relationships appears to have much in common with the distinctions drawn in GUM, substantial differences arise due to their very different purposes: i.e., SUMO does not try to cover the meaning of spatial language while GUM’s categorization is strictly based on the requirements of treating natural language semantics. Many cases motivated by linguistic examples are not then covered. Moreover, the term `SpatialRelation` does not seem to be used consistently in the SUMO categorization because it also reflects information about relations between size properties of entities, i.e., *larger* and *smaller*, which is covered by a `ScaledComparison` configuration in GUM and is independent of any specific spatial configuration. Most importantly, the central SUMO function `WhereFn` does not account at all for the complex construction of the relative position of an object and its relatum covered by GUM’s `GeneralizedLocation`.

Another ontology-like organization that is much more closely linked with linguistic evidence and which also includes some spatially relevant characterizations is the `FrameNet` lexical database [10], part of the `Berkeley FrameNet` project. Information within `FrameNet` is primarily motivated by lexicographic concerns and provides a representation of lexical semantics in terms of the ‘frames’ that words may appear in and the distinctive sets of ‘roles’ that such frames provide. Frames concerning spatial aspects include `Motion`, `Motion_scenario`, `Motion_directional`, `Moving_in_place`, `Cause_to_move_in_place` and `Quitting_a_place`. Each of these defines roles that express information such as the mover, the direction, the place, the path and so on. Although a fine-grained articulation of the semantics of spatial modalities and other particularly spatial details has not been one of `FrameNet`’s goals, it does exhibit several points of overlap with the distinctions drawn in GUM and a detailed investigation of alignment between the `FrameNet` examples and a GUM classification would be of considerable interest. Certain distinctions drawn in GUM—for example between placements and routes or between path placements and

²⁷ See also the corresponding formal specification at <http://www.ontospace.uni-bremen.de/ontology/gum-space/GUMwithDCC.html>.

²⁸ <http://research.cyc.com>.

path indications—do not appear to be reliably indicated in FrameNet, despite their distinct linguistic realizations. In such cases the FrameNet analyses appear to be *including contextualization* decisions in their assignment of roles. This is partly due to the lexical orientation of FrameNet, which, as we argued in Section 3 above, systematically undervalues the contribution of grammatical evidence. We would predict that this should impact negatively on automatic classification based on the FrameNet scheme.

Several more lexically-oriented classifications are similar in this respect, although nevertheless offering well-motivated groupings of lexical items that could beneficially be linked with the GUM-based lexical semantics. VerbNet [92],²⁹ for example, organizes verb entries by the range of grammatical alternations and argument restrictions that clauses containing the verbs enter into. This extends the basic work on alternations of Levin [109] and also exhibits similarities with our methodology of isolating grammatical syndromes. However, the approach is less inclusive in terms of the grammatical phenomena admitted as evidence and, perhaps as a result of this, the resulting organization is quite flat. Although a good source of information concerning the range of lexicalizations available, the distinction between different kinds of spatial relationships is not addressed in detail. Similar potentially useful but partial sources of linguistic information can be found in initiatives such as ‘The Preposition Project’ (TPP),³⁰ which contains senses of particular lexical prepositions, including those with spatial meaning.

In contrast to these directions, a non-lexically inspired ontology that specifically focuses on spatial aspects in natural language and which is intended to be used within a spatial application for human robot interaction is that mentioned above of Kruijff et al. [96]. This is closer to the organization proposed by GUM, although the grammatical evidence appealed to is less broad and there has been less of an attempt to evaluate the semantic representations against naturalistic data. The ontology does not then make the fine-grained distinctions between spatial relations and spatial configurations that can be found in the spatial extension of GUM. Moreover, the basic two-level semantics division between semantics and conceptual/contextual information is not maintained, leading to over-committed interpretations.

Finally, it is interesting to note that several approaches to processing spatial language now rely on linguistically-motivated classifications of the linguistic utterances that occur. However, these all provide a limited range of relevant abstractions due, on the one hand, to their task and application-specific focus and, on the other, to their appeal to a narrower range of linguistic evidence. Denis and colleagues [50,48], for example, provide a five-way classification of the components of spatial navigation, including actions, actions with landmarks, introduction of new landmarks, descriptions of landmarks and ‘others’. Similarly, Levit and Roy [112] simulate the interpretation of path descriptions (also with respect to the Map Task corpus that we mention below) by means of a probability-based combination of what they term *navigational information units* (NIUs). NIUs form an “intermediate” form of representation based on a subset of the literature of spatial linguistic expressions, and invoke categories such as TO (“in a straight line approach the closest point of a reference object”), AWAY_FROM (“move in the direction opposite to centre mass of a reference object”), PATH and so on. These categories form a small subset of those available within the Generalized Upper Model. However, although Levit and Roy’s system relied on hand-coded manual construction of NIUs—the connections between actual linguistic expressions and NIUs specifications had not been implemented—they argued nevertheless that such an “intermediate” representation is a valuable step in grounding certain spatial communication tasks, regardless of how it is achieved. The semantic specifications provided by GUM can be seen in precisely this light, providing a representation intermediate between utterances and contextualization. And, what is more, in the case of GUM the connection between linguistic form and semantic specification has been implemented. It would be an interesting further step, therefore, to examine whether the probabilistic modeling provided for the NIUs could also be extended to cover GUM specifications.

6. Ontology evaluation, coverage and agreement

While Section 5 discussed GUM’s adequacy and applicability with respect to compositionality, inference, and application, this section evaluates GUM with respect to data coverage and inter-annotator agreements when the ontology is used for linguistic data annotation. In both areas we will see that GUM appears to provide a significant and measurable advance in the treatment of spatial language and its semantics.

6.1. Broad corpus-driven evaluation of linguistic coverage

For the reasons explained above, we must consider the breadth of GUM’s coverage of linguistic data and not only its detail. To support this we have analyzed how GUM specifications fare when classifying natural language utterances taken from a broader set of test corpora. In addition, we attempt here to provide a further important aspect of quality control by providing evaluation metrics to benchmark coverage and to record gaps as well as growth as we move to new application areas.

The broader set of test corpora considered were deliberately selected in order to exhibit a considerable quantity of spatial language. For this we used seven different spatial language corpora containing natural language collected in English

²⁹ Cf. the ‘unified verb index’ available at: <http://verbs.colorado.edu/verb-index/index.php>.

³⁰ <http://www.cres.com/prepositions.html>.

Table 2

Distribution of spatial configuration types against analyzed corpora.

Configurations	MapTask CE1	Trains CE2	IBL CE3	Bielefeld CD1	AIBO CD2	Stuga CD3	Rolland CD4	Total
SpatialLocating	65	12	27	32	57	26	76	295
NonAffectingDirectedMotion	44	26	57	7	67	44	36	281
AffectingDirectedMotion	1	38		54	2		5	100
NonAffectingOrientationChange			10	1			11	22
NonAffectingOrienting			1	7		1		9
AffectingOrientationChange				3				3
NonAffectingMotion						1	1	2
Totals	110	76	95	104	126	72	129	712

and German. The use of spatial corpora rather than broader coverage resources such as the British National Corpus [29] was to ensure both that a breadth of spatial terminology was included and that complex constructions which involve a number of related spatial constructions were available in the data set. For English analysis, three spatial language corpora were selected: the Trains 93 Dialogues [72], the HCRC Map Task [3], and the IBL (Instruction Based Learning) Corpus [107]. For German, four spatial corpora were selected for analysis: the Bielefeld SFB 360 Corpus [141], and three corpora gathered in recent years within our own group: Aibo2 [56], Rolland [156,142], and Stuga (unpublished route description corpus).³¹ As all of these corpora are collections of task-oriented dialogues in the spatial domain, they contain a wide variety of spatial terms including localizing expressions, spatial action descriptions, and route descriptions—although the proportion of these features is by no means evenly distributed as will be discussed below.

Source data selection proceeded by collecting a sample set of “spatial language utterances” from each of the considered corpora. These utterances were selected on the basis of their containing either one or more clauses judged to convey spatial action, motion, or localization information directly, or a clause constituent which conveyed placement or direction information within another clause type. Clauses with idiomatic or metaphorical uses of spatial terms were not considered. For the TRAINS 93 corpus, four dialogues, i.e., D93-11.3, D93-15.5, D93-19.1, and D93-26.3, were selected at random, from which the first 100 utterances with spatial contributions were extracted. Similarly, the first 100 spatial utterances were extracted from three randomly selected IBL dialogues: u4, u15, and u17. For the Map Task dialogues, which are considerably longer and more complex than Trains or IBL dialogues, one dialogue was selected at random, q1nc4, from which the first 100 spatial utterances were similarly extracted. For the Bielefeld SFB 360 Corpus, 100 examples were taken from dialogues about assembling a propeller aircraft.³² Five examples each were selected from Dialog01 to Dialog19 together with five examples from Dialog20 to Dialog22 according to their variety of different spatial expressions. Similarly, 100 utterances were extracted from ID A001 to A051 of the Aibo2 Corpus, which mostly includes route instructions from human–robot interaction in a naive user exploration scenario. For the Rolland Corpus, 100 utterances from R001 to R034 were selected randomly, comprising utterances expressed in a further scene exploration task. For the Stuga Corpus, 30 utterances were extracted at random from D001 to D020, including expressions about route descriptions within a human–human–interaction scenario. The total test corpus consisted of 630 spatial utterances. Analysis then proceeded by ‘cleaning’ the surface language in each utterance to isolate the core surface spatial information conveyed. Specifically, non-spatial content such as conflated confirmation speech acts, conflated stalls or other communication management content, plus non-spatial modifications of nouns or verbs were removed so as to focus on spatial content. The resulting utterances do not constitute canonical forms, however, since they remain close to the original surface language with lexical choice and other structural information preserved for spatial content.

A GUM specification was then assigned for each pre-processed utterance, providing a ‘semantic annotation’ for each unit. This assignment was performed manually by native speakers of the analyzed language against the category descriptions provided by GUM and involved the bottom-up assignment of semantic categories to constituents of each utterance. The result was a complete GUM semantic description of the kind illustrated in previous sections. The corpus of 630 analyzed utterances gave 712 instances of spatially-relevant Configurations and 918 instances of SpatialModality. The utterances, their pre-processed forms, and the specifications of the surface spatial semantics are available online as part of the GUM documentation.³³ Of the total included in the target test corpus collection, a spatial semantics was successfully allocated to 592 of the utterances, leaving just over 6% of the corpus for which the spatial import of these utterances was not judged to be adequately covered. Segments for which no spatial semantics were judged adequate included a number of manner expressions and reciprocal spatial constructions; the coverage of these remains for future work.

This high degree of coverage is already a positive result for GUM. However, we can also use the statistics gathered both to isolate areas where further development should be undertaken and to pinpoint design requirements for further corpus collection activities. The GUM categories of Configuration and SpatialModality and their subconcepts provide a precise specification of a diverse range of linguistically construed spatial situations. When considering a corpus of spatial expressions,

³¹ This selection is due to the fact that the same breadth of public spatial corpora is not yet available for German as for English.

³² <http://www.sfb360.uni-bielefeld.de/transkript/b1-txt/>.

³³ See: <http://www.ontospace.uni-bremen.de/ontology/evaluation/gum-evaluation.html>.

Table 3

Distribution of spatial modality types against analyzed corpora ordered by total number of spatial modalities present.

SM	CE1	CE2	CE3	CD1	CD2	CD3	CD4	Total	SM	CE1	CE2	CE3	CD1	CD2	CD3	CD4	Total
GDN	8	81	22	3	53	7	8	182	Distr	2			3	8		1	14
LP	16		19	1	7	24	29	96	Cent	2			2	7		2	13
Sup	6		7	36	5	10	9	73	Card	1				11			12
RP	7		11	1	14	14	24	71	LPE					12			12
FP	3	1	4	2	11	13	10	44	RPE	4			2	4	1	1	12
Cont	1	14	1	12	2	3	7	40	BPE				4	5			9
PRE	8	1	13		4	3	9	38	DOFC		1	1	1			5	8
SD	19		12	2				33	API					8			8
FPE			2	2	15	11	3	33	QnD	2		1		1		2	6
GDD	2	20	3	1		1	3	30	BPI				3	2	1		6
UPE	15			14				29	Arc	3				1		2	6
QID	2			7		6	11	26	NPA							5	5
Prox	7	9	7			1	2	26	APE	5							5
OPE	8		4	14				26	PR			1			1	2	4
RNPA	5		14	2				21	AP				3				3
GD	7		2	5	3	1	3	21	Distal	1				2			3
Peri			8	1		10	1	20	RPI		2						2
BP	7	1		2	3	1	4	18	Topo	1							1
PRI			4	1	6	5	16	16	Seq					1			1
Sur	1		1		7	2	5	16	LPI					1			1

Abbreviations for the corpora as used above; abbreviations used for spatial modalities: AP: AboveProjection, APE: AboveProjectionExternal, API: AboveProjectionInternal, Arc: ArcDirectional, BP: BackProjection, BPE: BackProjectionExternal, BPI: BackProjectionInternal, Card: Cardinality, Cent: Central, Cont: Containment, DOFC: DenialOfFunctionalControl, Distal: Distal, Distr: Distribution, FCE: Functional Controlled External, FCI: FunctionalControlledInternal, FP: FrontProjection, FPE: FrontProjectionExternal, GD: GeneralDirectional, GDD: GeneralDirectionalDistancing, GDN: GeneralDirectionalNearing, LP: LeftProjection, LPE: LeftProjectionExternal, LPI: LeftProjectionInternal, NPA: NonProjectionAxial, OPE: OverProjectionExternal, Per: Peripheral, PR: PathRepresenting, PRE: PathRepresentingExternal, PRI: PathRepresentingInternal, Prox: Proximal, QnD: QuantitativeDistance, QtD: QualitativeDistance, RNPA: RelativeNonProjectionAxial, RP: RightProjection, RPE: RightProjectionExternal, RPI: RightProjectionInternal, SD: SpecificDirectional, Seq: Sequential, Sup: Support, Sur: Surrounding, Topo: TopographicDirectional, UPE: UnderProjectionExternal.

therefore, we can consider to what extent the diversity of the utterances occurring in the corpus overlap with the diversity of situations covered by the ontology. Our present results show that the spatial language corpora considered offer only a limited range of situations. If our randomly selected subset is in any way representative of their respective corpora (and we have no reason to presume it is not), then this lack of diversity limits our ability to effectively benchmark existing coverage. We see this concretely in the overviews of our results given in Tables 2 and 3, which show the distribution of spatial configuration types and spatial modality types against the corpora analyzed. In Table 2 it can be seen that the corpora contain a large number of spatial locating and directed motion configurations with a very much smaller number of orientation changes and orienting. In order to probe the accuracy and coverage of these other areas of GUM, it will be necessary to find many more naturally occurring examples—either in other corpora or by means of focused experiments which require expressions in the desired semantic areas.

Similarly, with the spatial modalities presented in Table 3, we see large proportions of uses for modalities such as GeneralDirectionalNearing, FrontProjection, LeftProjection, and Support, about ten further modalities having relatively frequent use, and several that are seen only occasionally in the data. Spatial modalities not seen at all in the data include abstract modalities which would be accounted for by sub-categories, e.g., Connection, and one terminal GUM spatial modality, MultipleDirectional, which is used to capture the non-directed motion in a use such as “they walked around town”. Again it would be desirable to find more examples of the infrequently occurring spatial modalities, as it is certain that there is much more detail to bring out here. The distribution that we observe also provides a semantic characterization of the particular text types that made up the originating corpora; the large proportion of route descriptions, for example, contribute significantly to the preponderance of general directionals, etc.

As the coverage of both the GUM linguistic ontology and the number of annotated corpora linked to the linguistic ontology grow, statistics of this kind will provide useful metadata concerning the material covered. This can then be used for guiding further development, exploring corpora containing lesser used spatial information or pulling apart already covered areas of spatial information in finer detail. We can also draw in other areas of spatial work and treat these as corpora for benchmarking GUM coverage: for example, several of the lexicographically-oriented lexicon projects mentioned above also include spatial examples relevant for benchmarking; similarly, considering linguistic annotations such as those investigated to date by the SpatialML spatial markup language initiative [116] would also usefully complement the range of linguistic phenomena considered to date.

6.2. Inter-annotator agreement

The GUM ontology can also be interpreted as an annotation schema for natural language sentences in its own right. This offers a further method for its evaluation, i.e., by carrying out annotation tasks and measuring the degree of agreement that different annotators achieve. We can then apply existing methods for evaluating the reliability and usability of the ontology

Table 4

Inter-annotator agreement: results for GUM's spatial configuration, relation, modality and modification of English data sample. The table shows pairwise comparisons between annotator E1, E2, and GS.

	Category	Percent agreement (%)	Cohen's Kappa	Agreements (No.)	Disagreements (No.)	Cases (No.)	Categories used (No.)
E1–E2	configuration	71.3	0.59	82	33	115	8
	spatial role	75.6	0.70	146	47	193	12
	modality	70.6	0.67	101	42	143	24
	modification	66.7	0.59	10	5	15	5
E1–GS	configuration	76.5	0.67	88	27	115	8
	spatial role	83.4	0.79	161	32	193	12
	modality	76.9	0.74	110	33	143	24
	modification	73.3	0.66	11	4	15	5
E2–GS	configuration	87.0	0.81	100	15	115	8
	spatial role	87.6	0.84	169	24	193	12
	modality	71.3	0.67	102	41	143	24
	modification	73.3	0.67	11	4	15	5

Table 5

Inter-annotator agreement: results for GUM's spatial configuration, relation, modality and modification of German data sample. The table shows pairwise comparisons between annotator G1, G2, and GS.

	Category	Percent agreement (%)	Cohen's Kappa	Agreements (No.)	Disagreements (No.)	Cases (No.)	Categories used (No.)
G1–G2	configuration	87.5	0.79	119	17	136	9
	spatial role	78.1	0.72	139	39	178	12
	modality	68.3	0.66	112	52	164	36
	modification	69.8	0.64	60	26	86	10
G1–GS	configuration	86.8	0.79	118	18	136	9
	spatial role	78.7	0.73	140	38	178	12
	modality	73.2	0.71	120	44	164	36
	modification	70.9	0.65	61	25	86	10
G2–GS	configuration	80.1	0.69	109	27	136	9
	spatial role	82.0	0.77	146	32	178	12
	modality	73.8	0.72	121	43	164	36
	modification	74.4	0.69	64	22	86	10

considered as an annotation scheme. Such measurements can be used to show the reliability and comprehensibility of GUM's specification: if different annotators annotate similar sentences in a similar way, this supports the claim that GUM provides regularly recognizable distinctions between categories. Inter-annotator agreement measures of this kind are not intended to prove that GUM represents actual human conceptual structures of space or the language of space; rather what can be shown is that GUM's categories can be learned by non-experts, that the categories (after a training phase) can be distinguished correctly, and that they are not randomly chosen for some linguistic terms.

Here we present results from an exploratory study of inter-annotator agreement that we have conducted with four annotators, two native German annotators (G1 and G2) and two native English annotators (E1 and E2), who were not previously familiar with GUM and its organization [79]. Based on the guidelines for organizing inter-annotator agreement studies set out by Lombard et al. [114], all annotators were provided with a manual on how to annotate sentences with GUM together with a spreadsheet document reflecting the structure of the GUM categories.³⁴ The annotation task was split into a training phase with 10 sentences, a supervised annotation phase with 2×50 sentences, and an unsupervised annotation sample with 100 German and 90 English sentences. Inter-annotator agreement was calculated for the unsupervised annotation samples. All of these sentences were taken from the data set used to measure GUM's coverage (cf. Section 6.1). The GUM specifications from this data set were used as a "gold standard" [69], counting them as a 'third annotator' (GS) per language. Annotators were instructed (i) to clean up the sentences, i.e., to remove non-spatial information from the sentences and grammatically correct them if necessary, and (ii) to annotate the sentences according to the GUM specification as given by the spreadsheet document.

We compare the annotation results based on the different parts of the GUM specification, namely the particular subtypes of Configuration (such as SpatialLocating, NonAffectingDirectedMotion, etc.), attribute (such as placement, direction, etc.), SpatialModality (such as LeftProjectionExternal, Proximal, etc.) and modification (such as extension, accessibility, etc.) that the annotators selected. Hence, the major decision points within the annotation or specification for a sentence were investigated. The results of agreements between the annotators E1, E2, and GS for the English data sample are shown in Table 4; results for G1, G2, and GS for the German data sample are shown in Table 5.

³⁴ The annotation sources are available at <http://www.ontospace.uni-bremen.de/ontology/evaluation/gum-evaluation.html>.

The English sample with 90 sentences, which were taken from Trains and IBL (see Section 6.1), contains 115 configurations, 193 spatial roles, 143 modalities, and 15 modifications that were used for comparing agreement. Due to the clean-up task, we had to remove 15 additional configurations as the annotations did not refer to the same sentences. The calculation of agreement was performed in a ‘strict’ fashion, i.e., even similar categories (e.g., *LeftProjection* and *LeftProjectionInternal*, or *direction* and *motionDirection*) were regarded here as different annotations and marked as disagreements. The results in Table 4 show agreement on all categories to be above 70%, except for the agreement on modifications between annotator E1 and E2. Modifications, however, only occurred in 15 cases. Particularly promising are the results for spatial modalities: a wide range of 24 different categories were used by the annotators and their agreement is still higher than 70%. Also Cohen’s Kappa, a standard measure of agreement [41], is on average slightly above 0.7 over all comparisons, which is also a good positive indication of the reliability of the annotation decisions made.

The German sample with 100 sentences was taken from Rolland and Aibo (see Section 6.1) and contains 136 configurations, 178 spatial roles, 164 modalities, and 86 modifications. Agreements between annotators were compared as for the English annotations. Due to the clean-up task, 16 additional configurations were removed, and as with the English sample, the calculation of agreement was performed ‘strictly’. The results in Table 5 show agreement on all categories above 70%, except for the agreement on modalities and modifications between annotator G1 and G2. The number of modalities that were used, however, is 36 and this therefore shows a broad variety of different types. If we would factor out over- or underspecifications (for instance, combining *FrontProjectionExternal* with *FrontProjection*, etc.), the agreement is above 80%; we omit such calculation here, however. Also Cohen’s Kappa is in most cases above 0.7 for all comparisons, which again indicates reliability similar to the English corpus samples.

An interesting finding from the annotation evaluation is that some annotations illustrate dependencies from the modality hierarchy (cf. Fig. 7)—even without explicit knowledge of the annotators since they were not informed about the modalities’ hierarchical structure. In sentence (41) taken from the German sample, for instance, the annotation of the ‘gold standard’ for *gegenüber* (*opposite side*) is the modality *Proximal*. Both German annotators, however, annotated this modality as a *FrontProjectionExternal*. Although this category is too specific for the relationship of being on the opposite side, it is a subcategory of *Proximal*. Hence, both categories shows a strong connection—formalized by the hierarchical relationship in GUM and implicitly indicated by the annotations.

(41) Es ist gegenüber von mir. [It is opposite me.]

On average the results show a high agreement between the different annotators. This also proves that with the use of a manual, uninformed annotators can be taught to apply the decision options in GUM and to decide on a specific configuration with its spatial roles filled with a high variety of different modality types and modifications.

6.3. Discussion

The empirical, corpus-driven approach to ontology evaluation presented in this section has been useful for several reasons. First, we have seen both that the GUM spatial extension covers existing spatial language corpora well and that those corpora need to be extended in breadth to provide more data for further linguistic ontology development. And second, we have demonstrated both that it is possible to use a linguistic ontology to inform the semantic annotation of linguistic data and that, when this is done, the results show a promising degree of reliability. All of these directions need to be taken further in future research.

7. Conclusion and outlook

Our starting point in this paper was the extremely flexible relationship observed between spatial language and contextualized interpretations of that language. There is an urgent need for versatile and comprehensive accounts that support the contextualization process by pinpointing the information that needs to be anchored by context without prematurely over-committing to particular spatial interpretations. We have addressed this by considering the *linguistic* construction of space far more closely than has hitherto been the case. This has led to a ‘linguistically responsible’ characterization of the semantic distinctions that are carried by grammar (at least in English and German), couched as an extension of the Generalized Upper Model linguistic ontology and employing current ontological engineering principles.

We have shown that this treatment covers a large proportion of naturally occurring linguistic expressions involving space. Simple inventories of linguistic terms (and their direct semantic interpretation in terms of physical spatial models) have been replaced by a richly structured characterization of linguistically-motivated spatial semantics that provides strong support for active mediation between linguistic form and detailed spatial models. This characterization generalizes across contexts of use and applications, just as the corresponding linguistic expressions do. The Generalized Upper Model spatial extension is now accordingly employed as a level of linguistic semantics for both automatic generation and analysis within spatially-aware computational systems. We have seen that GUM lies within the description logic *ALCHQ(D)* and, as a consequence, any appropriate DL reasoner can be applied to provide reasoning support. Moreover, to support usability further, GUM is accessible online as an OWL DL specification including detailed comments about its categories and relations

as documentation.³⁵ At a design level, the OntoClean methodology has been applied throughout GUM's development, strictly enforcing ontological requirements of consistency and respect for identity criteria.

We claim that an adequate account of linguistic spatial expressions will require at least the kinds of distinctions that our characterization has set out, independently of how these distinctions are then anchored in axiomatizations of space of particular kinds, in action routines for embodied behavior, or in perceptual models. The linguistic variability deployed in spatial expressions exhibits the distinctions drawn and so brings constraints to bear on the properties that any such models need to provide. The categories of the formalization we develop thereby support mediation, by virtue of the constraints they bring, but without imposing commitment to particular underlying formal systems. This allows us to remain relatively agnostic about the spatial models employed in any particular application system or theoretical account, thereby further supporting re-use.

We are now extending the coverage of GUM, benchmarking progress against a broadening range of corpora of naturally occurring spatial language. Particular areas being added include that required for describing geographic information, as explored for the spatial language annotation task within SpatialML [116], and the consequences of pursuing similar descriptions in other natural languages. Setting out explicit connections between the categories of GUM and the semantic requirements of languages other than English and German will be crucial for determining the extent to which the spatial extension of GUM needs to vary across languages. In this area also, there is a substantial body of previous work to draw on [6,32,76,153]; at present we assume both that further spatial modalities will be discovered and that there will be variations in the spatial modalities required by specific languages. Moreover, in all cases of extension, we will continue to develop the formalization of the spatial interpretation, using GUM as the linguistic component of a formally specified heterogeneous account mediating between language use and contextualized interpretation.

Finally, the account is now sufficiently detailed and concrete, that we can consider how to extend the range of evidence that can be drawn upon for investigating its adequacy and appropriateness. We have described our use and evaluation of the GUM organization as an annotation scheme above, and this needs to be taken further. We can also investigate in greater depth the formal relationships between the spatial theories that are taken as appropriate interpretations of distinct portions of the GUM concept hierarchy—here it will be interesting to appeal to current research exploring characterizations of spatial calculi in general in order to see if parallels between the GUM organization and the formal 'meta'-structure of this domain can be isolated. There is also the possibility of psychological investigation: now that a firm relationship has been posited between distinct areas of spatial linguistic semantics and linguistic forms, we can explore to what extent this is supported empirically. Carlson and colleagues (e.g., [35,34]), for example, have performed experiments which show challenging and 'non-standard' connections between the spatial information activated by particular linguistic terms, such as 'distance' information being activated by projective terms such as 'left'/'right', etc. Investigations of whether such activations follow the connections suggested within GUM may throw new light on its organization as a whole. Moreover, as Carlson and van Deman conclude:

"The premise in the literature has been that different terms are associated with different information. ... However, the current results indicate that a distinction must be made between the semantics conveyed by the term itself, and the information deemed relevant during the processing of the spatial term" [35, p. 434].

We see this as offering potential support for precisely the kind of two-level architecture that we have proposed: the linguistic ontology provides the semantics of the 'terms themselves', but when this is anchored to contextualized spatial interpretations further kinds of information may naturally be entailed. There is then considerable opportunity for targeted investigation using both the general architecture we propose and the particular distinctions we have developed within this.

References

- [1] M. Aiello, I. Pratt-Hartmann, J.v. Benthem (Eds.), *Handbook of Spatial Logics*, Springer-Verlag, Berlin, 2007.
- [2] H. Alshawi (Ed.), *The Core Language Engine*, MIT Press, Cambridge, MA, 1992.
- [3] A.H. Anderson, M. Bader, E. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. Thompson, R. Weinert, The HCRC map task corpus, *Language and Speech* 34 (4) (1991) 351–366.
- [4] G. Antoniou, A. Kehagias, A note on the refinement of ontologies, *International Journal of Intelligent Systems* 15 (2000) 623–632.
- [5] N. Asher, Objects, locations and complex types, in: M. Aurnague, M. Hickmann, L. Vieu (Eds.), *The Categorization of Spatial Entities in Language and Cognition*, in: *Human Cognitive Processing*, vol. 20, John Benjamins Publishing Company, Amsterdam, The Netherlands, 2007, pp. 337–361.
- [6] N. Asher, P. Sablayrolles, A typology and discourse semantics for motion verbs and spatial PPs in French, *Journal of Semantics* 12 (1) (1995) 163–209.
- [7] M. Aurnague, M. Hickmann, L. Vieu (Eds.), *The Categorization of Spatial Entities in Language and Cognition*, Human Cognitive Processing, vol. 20, John Benjamins Publishing Company, Amsterdam, The Netherlands, 2007.
- [8] M. Aurnague, L. Vieu, A three-level approach to the semantics of space, in: C. Zelinski-Wibbelt (Ed.), *Semantics of Prepositions: From Mental Processing to Natural Language Processing*, in: *Natural Language Processing*, vol. 3, Mouton de Gruyter, Berlin, 1993, pp. 393–439.
- [9] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, P. Patel-Schneider (Eds.), *The Description Logic Handbook: Theory, Implementation and Applications*, Cambridge University Press, Cambridge, UK, 2003.
- [10] C.F. Baker, C.J. Fillmore, J.B. Lowe, The Berkeley FrameNet Project, in: *Proceedings of the ACL/COLING-98, 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, Association for Computational Linguistics and Morgan Kaufmann Publishers, San Mateo, CA, 1998.

³⁵ <http://www.ontospace.uni-bremen.de/ontology/stable/GUM-3-space.owl>.

- [11] J. Baldridge, G.-J. Kruijff, Coupling CCG and hybrid logic dependency semantics, in: Proceedings of 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, Pennsylvania, 2002.
- [12] J.A. Bateman, Upper modeling: Organizing knowledge for natural language processing, in: Proceedings of the Fifth International Natural Language Generation Workshop, Pittsburgh, PA, 1990, organized by Kathleen R. McKeown (Columbia University), Johanna D. Moore (University of Pittsburgh) and Sergei Nirenburg (Carnegie Mellon University). Held 3–6 June 1990, Dawson, PA. URL: <http://acl.ldc.upenn.edu/W/W90/W90-0108.pdf>.
- [13] J.A. Bateman, Language and space: A two-level semantic approach based on principles of ontological engineering, *International Journal of Speech Technology* 13 (1) (2010) 29–48.
- [14] J.A. Bateman, S. Borgo, K. Lüttich, C. Masolo, T. Mossakowski, Ontological modularity and spatial diversity, *Spatial Cognition and Computation* 7 (1) (2007) 97–128.
- [15] J.A. Bateman, S. Farrar, Spatial ontology baseline, SFB/TR8 internal report I1-[OntoSpace]: D2, Collaborative Research Center for Spatial Cognition, University of Bremen, Germany, 2004.
- [16] J.A. Bateman, R.T. Kasper, J.D. Moore, R.A. Whitney, A general organization of Knowledge for natural language processing: The PENMAN Upper Model, Tech. rep., USC/Information Sciences Institute, Marina del Rey, California, 1990.
- [17] J.A. Bateman, I. Kruijff-Korbayová, G.-J. Kruijff, Multilingual resource sharing across both related and unrelated languages: An implemented, open-source framework for practical natural language generation, *Research on Language and Computation* 3 (2) (2005) 191–219.
- [18] J.A. Bateman, B. Magnini, G. Fabris, The generalized upper model knowledge base: Organization and use, in: N.J.I. Mars (Ed.), *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*, IOS Press, Amsterdam, 1995.
- [19] J.A. Bateman, T. Tenbrink, S. Farrar, The role of conceptual and linguistic ontologies in discourse, *Discourse Processes* 44 (3) (2007) 175–213.
- [20] C. Becker, D. Nicklas, Where do spatial context-models end and where do ontologies start? A proposal of a combined approach, in: J. Indulska, D.D. Roure (Eds.), *Proceedings of the First International Workshop on Advanced Context Modelling, Reasoning and Management in Conjunction with UbiComp 2004*, University of Southampton, 2004.
- [21] B. Bennett, The application of qualitative spatial reasoning to GIS, in: R. Abraham (Ed.), *Proceedings of the First International Conference on GeoComputation*, vol. I, Leeds, 1996.
- [22] B. Bennett, Space, time, matter and things, in: C. Welty, B. Smith (Eds.), *Formal Ontology in Information Systems: Proceedings of the 2nd International Conference on Formal Ontology in Information Systems*, ACM Press, New York, 2001.
- [23] B. Bennett, What is a forest? On the vagueness of certain geographic concepts, *Topoi* 20 (2) (2001) 189–201.
- [24] B. Bennett, P. Agarwal, Semantic categories underlying the meaning of 'Place', in: S. Winter, M. Duckham, L. Kulik, B. Kuipers (Eds.), *Spatial Information Theory: Proceedings of the 8th International Conference, COSIT-2007, Melbourne, Australia*, in: *Lecture Notes in Computer Science*, vol. 4736, Springer-Verlag, Berlin, Heidelberg, 2007.
- [25] M. Bierwisch, On the grammar of local prepositions, in: M. Bierwisch, W. Motsch, I. Zimmermann (Eds.), *Syntax, Semantik, und Lexikon: Rudolf Ruzicka zum 65. Geburtstag*, Akademie Verlag, Berlin, 1988, pp. 1–65.
- [26] M. Bierwisch, How much space gets into language, in: P. Bloom, M.A. Peterson, L. Nadel, M.F. Garrett (Eds.), *Language and Space*, MIT Press, Cambridge, MA, 1999, pp. 31–76.
- [27] M. Bierwisch, E. Lang (Eds.), *Dimensional Adjectives, Grammatical Structure and Conceptual Interpretation*, Springer, Berlin, 1989.
- [28] P. Bloom, M. Peterson, L. Nadel, M. Garrett (Eds.), *Language and Space*, MIT Press, Cambridge, MA, 1996.
- [29] BNC Consortium, *The British National Corpus, version 3 (BNC XML Edition)*, distributed by Oxford University Computing Services on behalf of the BNC Consortium, <http://www.natcorp.ox.ac.uk/>, 2007.
- [30] S. Borgo, N. Guarino, C. Masolo, A pointless theory of space based on strong congruence and connection, in: L. Aiello, J. Doyle (Eds.), *Principles of Knowledge Representation and Reasoning, Proceedings of the 5th International Conference KR96*, Morgan Kaufmann, San Mateo, CA, 1996.
- [31] S. Borgo, N. Guarino, C. Masolo, Stratified ontologies: The case of physical objects, in: *Proceedings of the Workshop on Ontological Engineering at ECAI'96*, Budapest, Hungary, 1996.
- [32] M. Bowerman, Learning how to structure space for language: A crosslinguistic perspective, in: P. Bloom, M.A. Peterson, L. Nadel, M.F. Garrett (Eds.), *Language and Space*, MIT Press, Cambridge, MA, 1999, pp. 385–436.
- [33] H. Buxton, R.J. Howarth, Generating dynamic scene descriptions, in: P. Olivier, K.-P. Gapp (Eds.), *Representation and Processing of Spatial Expressions*, Lawrence Erlbaum Associates, Mahwah, NJ, 1998, pp. 91–102.
- [34] L.A. Carlson, E.S. Covey, How far is near? Inferring distance from spatial descriptions, *Language and Cognitive Processes* 20 (5) (2005) 617–631.
- [35] L.A. Carlson, S.R. Van Deman, The space in spatial language, *Journal of Memory and Language* 51 (2004) 418–436.
- [36] L.A. Carlson, E. van der Zee (Eds.), *Functional Features in Language and Space: Insights from Perception, Categorization and Development*, Oxford University Press, Oxford, 2005.
- [37] L.A. Carlson-Radvansky, G.D. Logan, The influence of reference frame selection on spatial template construction, *Journal of Memory and Language* 37 (1997) 411–437.
- [38] R. Casati, B. Smith, A.C. Varzi, Ontological tools for geographic representation, in: N. Guarino (Ed.), *Formal Ontology in Information Systems (FOIS)*, IOS Press, Amsterdam, 1998, pp. 77–85.
- [39] R. Casati, A.C. Varzi, *Parts and Places: The Structures of Spatial Representation*, MIT Press (Bradford Books), Cambridge, MA and London, 1999.
- [40] P. Cimiano, U. Reyle, Towards foundational semantics—Ontological semantics revisited, in: B. Bennett, C. Fellbaum (Eds.), *Proceedings of the International Conference on Formal Ontology in Information Systems (FOIS)*, IOS Press, 2006. URL: <http://www.aifb.uni-karlsruhe.de/WBS/pci/Publications/fois06.pdf>.
- [41] J. Cohen, A coefficient of agreement for nominal scales, *Educational and Psychological Measurement* 20 (1) (1960) 37–46.
- [42] A. Cohn, S. Hazarika, Qualitative spatial representation and reasoning: An overview, *Fundamenta Informaticae* 43 (2001) 2–32.
- [43] A.G. Cohn, B. Bennett, J. Gooday, N.M. Gotts, Representing and reasoning with qualitative spatial relations, in: O. Stock (Ed.), *Spatial and Temporal Reasoning*, Kluwer Academic Publishers, Dordrecht, 1997, pp. 97–132.
- [44] K. Coventry, T. Tenbrink, J. Bateman (Eds.), *Spatial Language and Dialogue*, Oxford University Press, Oxford, 2009.
- [45] K.R. Coventry, S.C. Garrod, Saying, Seeing and Acting, *The Psychological Semantics of Spatial Prepositions*, Essays in Cognitive Psychology Series, Psychology Press, Hove, UK, 2004.
- [46] K.R. Coventry, M. Prat-Sala, L.V. Richards, The interplay between geometry and function in the comprehension of 'over', 'under', 'above' and 'below', *Journal of Memory and Language* 44 (2001) 376–398.
- [47] L.G. Creary, J.M. Gawron, J. Nerbonne, Reference to locations, in: *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Vancouver, British Columbia, Canada, 1989. URL: <http://www.aclweb.org/anthology/P89-1006>.
- [48] M. Daniel, M. Denis, Spatial descriptions as navigational aids: A cognitive analysis of route directions, *Kognitionswissenschaft* 7 (1998) 45–52.
- [49] D. Davidson, The logical form of action sentences, in: N. Rescher (Ed.), *The Logic of Decision and Action*, University of Pittsburgh Press, Pittsburgh, Pennsylvania, 1967, pp. 81–95.
- [50] M. Denis, The description of routes: A cognitive approach to the production of spatial discourse, *Cahiers de Psychologie Cognitive* 16 (1997) 409–458.
- [51] M.O. Dzikovksa, J.F. Allen, M.D. Swift, Linking semantic and knowledge representations in a multi-domain dialogue system, *Journal of Logic and Computation* 18 (3) (2008) 405–430.

- [52] M.J. Egenhofer, Towards the semantic geospatial web, in: Proceedings of the Geographic Information Science Conference (GIS'02), ACM, McLean, Virginia, 2002.
- [53] C. Eschenbach, Geometric structures of frames of reference and natural language semantics, *Spatial Cognition and Computation* 1 (4) (1999) 329–348.
- [54] C. Eschenbach, Contextual, functional, and geometric components in the semantics of projective terms, in: L. Carlson, E. van der Zee (Eds.), *Functional Features in Language and Space: Insights from Perception, Categorization, and Development*, Oxford University Press, Oxford, 2005, pp. 71–91.
- [55] M. Feist, D. Gentner, On plates, bowls, and dishes: Factors in the use of English 'in' and 'on', in: Proceedings of the 20th Annual Conference of the Cognitive Science Society, 1998.
- [56] K. Fischer, The role of users' concepts of the robot in human–robot spatial instruction, in: T. Barkowsky, M. Knauff, G. Ligozat, D. Montello (Eds.), *Spatial Cognition V: Reasoning, Action, Interaction*, Springer, Berlin, Heidelberg, 2007, pp. 76–89.
- [57] F.T. Fonseca, M.J. Egenhofer, P. Agouris, G. Câmara, Using ontologies for integrated geographic information systems, *Transactions in GIS* 6 (3).
- [58] N. Francez, M. Steedman, Categorical grammar and the semantics of contextual prepositional phrases, *Linguistics and Philosophy* 29 (2006) 381–417.
- [59] A.U. Frank, Spatial ontology: A geographical information point of view, in: O. Stock (Ed.), *Spatial and Temporal Reasoning*, Kluwer Academic Publishers, Dordrecht, 1997, pp. 135–153.
- [60] A.U. Frank, Tiers of ontology and consistency constraints in geographical information systems, *International Journal of Geographical Information Science* 15 (7) (2001) 667–678.
- [61] W. Frawley, *Linguistic Semantics*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1992.
- [62] C. Freksa, Using orientation information for qualitative spatial reasoning, in: A.U. Frank, I. Campari, U. Formentini (Eds.), *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, in: LNCS, vol. 639, Springer, Berlin, 1992, pp. 162–178.
- [63] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, Sweetening ontologies with DOLCE, *AI Magazine* 24 (3) (2003) 13–24.
- [64] K. Gapp, Basic meanings of spatial relations: Computation and evaluation in 3D space, in: Proceedings of AAAI-94, Seattle, Washington, 1994.
- [65] N. Guarino, The ontological level, in: R. Casati, B. Smith, G. White (Eds.), *Philosophy and the Cognitive Sciences*, Hölder–Pichler–Tempsky, Vienna, 1994, pp. 443–456.
- [66] N. Guarino, Formal ontology, conceptual analysis and knowledge representation, *International Journal of Human–Computer Studies* 43 (5/6) (1995) 625–640.
- [67] N. Guarino, Formal ontology and information systems, in: N. Guarino (Ed.), *Formal Ontology in Information Systems (FOIS)*, IOS Press, Amsterdam, 1998, pp. 3–18.
- [68] N. Guarino, C. Welty, An overview of OntoClean, in: S. Staab, R. Studer (Eds.), *Handbook on Ontologies*, Springer-Verlag, Heidelberg and Berlin, 2004, pp. 151–171.
- [69] K. Gwet, *Handbook of Inter-Rater Reliability*, Statax Publishing Company, 2001.
- [70] M.A.K. Halliday, C.M.I.M. Matthiessen, *Construing Experience Through Meaning: A Language-Based Approach to Cognition*, Cassell, London, 1999.
- [71] M.A.K. Halliday, C.M.I.M. Matthiessen, *An Introduction to Functional Grammar*, 3rd ed., Edward Arnold, London, 2004.
- [72] P.A. Heeman, J. Allen, The Trains 93 Dialogues, *Trains Technical Note 94-2*, Computer Science Dept., University of Rochester (Mar 1995). URL ftp://ftp.cs.rochester.edu/pub/papers/ai/95.tn2.Trains_93_dialogues.ps.gz.
- [73] T. Herrmann, J. Grabowski, *Sprechen: Psychologie der Sprachproduktion*, Spektrum Verlag, Heidelberg, 1994.
- [74] A. Herskovits, *Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions in English*, Studies in Natural Language Processing, Cambridge University Press, London, 1986.
- [75] M. Herweg, Aspectual requirements of temporal connectives: Evidence for a two-level approach to semantics, in: J. Pustejovsky, S. Bergler (Eds.), *Proceedings of the 1991 ACL Workshop on Lexical Semantics and Knowledge Representation*, Berkeley, CA, 1991.
- [76] M. Hickmann, Static and dynamic location in French: Developmental and cross-linguistic perspectives, in: M. Aurnague, M. Hickmann, L. Vieu (Eds.), *The Categorization of Spatial Entities in Language and Cognition*, in: *Human Cognitive Processing*, vol. 20, John Benjamins Publishing Company, Amsterdam, The Netherlands, 2007, pp. 205–232.
- [77] M. Hickmann, S. Robert (Eds.), *Space in Languages: Linguistic Systems and Cognitive Categories*, Typological Studies in Language, vol. 66, John Benjamins, Amsterdam, 2006.
- [78] J.R. Hobbs, Sketch of an ontology underlying the way we talk about the world, *International Journal of Human–Computer Studies* 43 (5/6) (1995) 819–830.
- [79] J. Hois, Inter-annotator agreement on a linguistic ontology for spatial language—A case study for GUM-Space, in: Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010), LREC, 2010.
- [80] J. Hois, O. Kutz, Counterparts in language and space-similarity and \mathcal{S} -connection, in: C. Eschenbach, M. Grüninger (Eds.), *Proceedings of the International Conference on Formal Ontology in Information Systems (FOIS)*, IOS Press, Amsterdam, 2008.
- [81] J. Hois, O. Kutz, Natural language meets spatial calculi, in: C. Freksa, N.S. Newcombe, P. Gärdenfors, S. Wöflf (Eds.), *Spatial Cognition VI: Learning, Reasoning and Talking about Space*, International Conference, Spatial Cognition 2008, Freiburg, Germany, in: *Lecture Notes in Artificial Intelligence*, vol. 5241, Springer, 2008.
- [82] J. Hois, O. Kutz, J.A. Bateman, Similarity-connections between natural language and spatial situations, in: K. Coventry, R. O'Ceallaigh (Eds.), *Workshop on Spatial Language in Context: Computational and Theoretical Approaches to Situation Specific Meaning (in Association with Spatial Cognition, 2008)*, 2008.
- [83] J. Hois, T. Tenbrink, R. Ross, J. Bateman, GUM-Space—The Generalized Upper Model spatial extension: A linguistically-motivated ontology for the semantics of spatial language, version 3.0, SFB/TR8 Technical report, Collaborative Research Center for Spatial Cognition, University of Bremen, Germany, 2009.
- [84] E.H. Hovy, Using an ontology to simplify data access, *Communications of the ACM* 46 (1) (2003) 47–49. URL: <http://doi.acm.org/10.1145/602421.602447>.
- [85] C. Huang, N. Calzolari, A. Gangemi, A. Lenzi, A. Oltramari, L. Prevot (Eds.), *Ontologies and the Lexicon*, Cambridge University Press, Cambridge, 2008.
- [86] R. Jackendoff, *Semantics and Cognition*, The MIT Press, Cambridge, MA, 1983.
- [87] R. Jackendoff, The architecture of the linguistic–spatial interface, in: P. Bloom, M.A. Peterson, L. Nadel, M.F. Garrett (Eds.), *Language and Space*, MIT Press, Cambridge, MA, 1999, pp. 1–30.
- [88] F. Jørgensen, J.T. Lønning, A minimal recursion semantic analysis of locatives, *Computational Linguistics* 35 (2) (2009) 229–270. URL: <http://aclweb.org/anthology-new/J/J09/J09-2004.pdf>.
- [89] Y. Kalfoglou, M. Schorlemmer, Ontology mapping: The state of the art, *Knowledge Engineering Review* 18 (1) (2003) 1–31. URL: <http://citeseer.ist.psu.edu/kalfoglou03ontology.html>.
- [90] J. Kalita, B. Badler, Interpreting prepositions physically, in: Proceedings of AAAI-91, 1991.
- [91] J.D. Kelleher, F.J. Costello, Applying computational models of spatial prepositions to visually situated dialog, *Computational Linguistics* 35 (2) (2009) 119–149. URL: <http://aclweb.org/anthology-new/J/J09/J09-2005.pdf>.
- [92] K. Kipper-Schuler, *VerbNet: A broad-coverage, comprehensive verb lexicon*, Ph.D. thesis, Computer and Information Science Department, University of Pennsylvania, Philadelphia, PA, 2005. URL: <http://repository.upenn.edu/dissertations/AAI3179808/>.
- [93] A. Klippel, T. Tenbrink, D. Montello, The role of structure and function in the conceptualization of directions, in: M. Dimitrova-Vulchanova, E. van der Zee (Eds.), *Motion Encoding in Spatial Language*, Oxford University Press, Oxford, forthcoming.

- [94] M. Kracht, On the semantics of locatives, *Linguistics and Philosophy* 25 (2002) 157–232.
- [95] M. Kracht, The fine structure of spatial expressions, in: A. Asbury, J. Dotlacil, B. Gehrke, R. Nouwen (Eds.), *The Structure of Local P*, John Benjamins, Amsterdam, 2008, pp. 35–62.
- [96] G.-J.M. Kruijff, H. Zender, P. Jensfelt, H.I. Christensen, Situated dialogue and spatial organization: What, where... and why? *International Journal of Advanced Robotic Systems* 4 (1) (2007) 125–138 (Special Issue on Human–Robot Interaction).
- [97] W. Kuhn, An image-schematic account of spatial categories, in: *Spatial Information Theory*, 8th International Conference, COSIT 2007, Melbourne, Australia, in: *Lecture Notes in Computer Science*, vol. 4736, Springer-Verlag, 2007, pp. 152–168.
- [98] O. Kutz, J. Hois, J. Bao, B. Cuenca Grau (Eds.), *Modular Ontologies: Proceedings of the 4th International Workshop (WoMo 2010)*, *Frontiers in Artificial Intelligence and Applications*, vol. 210, IOS Press, Amsterdam, The Netherlands, 2010.
- [99] O. Kutz, D. Lücke, T. Mossakowski, Heterogeneously structured ontologies—Integration, connection, and refinement, in: T. Meyer, M.A. Orgun (Eds.), *Advances in Ontologies. Proceedings of the Knowledge Representation Ontology Workshop (KROW 2008)*, in: *CRPIT*, vol. 90, ACS, Sydney, Australia, 2008.
- [100] O. Kutz, C. Lutz, F. Wolter, M. Zakharyashev, \mathcal{E} -Connections of abstract description systems, *Artificial Intelligence* 156 (1) (2004) 1–73.
- [101] O. Kutz, T. Mossakowski, Conservativity in structured ontologies, in: M. Ghallab, C. Spyropoulos, N. Fakotakis, N. Avouris (Eds.), *Proceedings of the 18th European Conference on Artificial Intelligence (ECAI-08)*, Patras, Greece, in: *Frontiers in Artificial Intelligence and Applications*, vol. 178, IOS Press, 2008. URL: <http://www.informatik.uni-bremen.de/~till/papers/modules-ECAI.pdf>.
- [102] O. Kutz, T. Mossakowski, M. Codescu, Shapes of alignments: Construction, combination, and computation, in: U. Sattler, A. Tamilin (Eds.), *Proceedings of the 1st Workshop on Ontologies: Reasoning and Modularity (WORM-08)*, No. 348, CEUR-WS, ESWC, Tenerife, Spain, 2008.
- [103] B. Landau, R. Jackendoff, 'What' and 'where' in spatial language and cognition, *Behavioral and Brain Sciences* 16 (1993) 217–238.
- [104] E. Lang, A two-level approach to projective prepositions, in: G. Rauh (Ed.), *Approaches to Prepositions*, Gunter Narr, Tübingen, 1991, pp. 127–167.
- [105] E. Lang, K.-U. Carstensen, G. Simmons, *Modelling Spatial Knowledge on a Linguistic Basis: Theory–Prototype–Integration*, *Lecture Notes in Artificial Intelligence*, vol. 481, Springer, Berlin, Heidelberg, 1991.
- [106] R.W. Langacker, *Foundations in Cognitive Grammar*, vol. 1, Theoretical Prerequisites, Stanford University Press, Stanford, CA, 1987.
- [107] S. Lauria, G. Bugmann, T. Kyriacou, J. Bos, E. Klein, Training personal robots via natural-language instructions, *IEEE Intelligent Systems* (2001) 38–45.
- [108] W. Levelt, Perspective taking and ellipsis in spatial descriptions, in: P. Bloom, M. Peterson, L. Nadel, M. Garrett (Eds.), *Language and Space*, MIT Press, Cambridge, MA, 1996, pp. 77–109.
- [109] B. Levin, *English Verb Classes and Alternations: A Preliminary Investigation*, University of Chicago Press, Chicago and London, 1993.
- [110] S.C. Levinson, *Space in Language and Cognition: Explorations in Cognitive Diversity*, Cambridge University Press, Cambridge, 2003.
- [111] S.C. Levinson, S. Kita, D.B. Haun, B.H. Rasch, Returning the tables: Language affects spatial reasoning, *Cognition* 84 (2002) 155–188.
- [112] M. Levit, D. Roy, Interpretation of spatial language in a map navigation task, *IEEE Transactions on Systems, Man and Cybernetics* 37 (3) (2006) 667–679.
- [113] G.D. Logan, D.D. Sadler, A computational analysis of the apprehension of spatial relations, in: P. Bloom, M.A. Peterson, L. Nadel, M.F. Garrett (Eds.), *Language and Space*, MIT Press, Cambridge, MA, 1999, pp. 493–530.
- [114] M. Lombard, J. Snyder-Duch, C.C. Bracken, Content analysis in mass communication: Assessment and reporting of intercoder reliability, *Human Communication Research* 28 (4) (2002) 587–604.
- [115] R.M. MacGregor, Inside the LOOM description classifier, *ACM SIGART Bulletin* 2 (3) (1991) 88–92.
- [116] I. Mani, J. Hitzeman, J. Richer, D.R. Harris, R. Quimby, B. Wellner, *SpatialML: Annotation scheme, corpora, and tools*, in: *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, 2008. URL: <http://sourceforge.net/projects/spatialml>.
- [117] W.C. Mann, Y. Arens, C.M.I.M. Matthiessen, S. Naberschnig, N.K. Sondheimer, *Janus abstraction structure—draft 2*, Technical report, USC/Information Sciences Institute, Marina del Rey, California, Oct. 1985 (circulated in draft form only).
- [118] D. Mark, W. Kuhn, B. Smith, A. Turk, *Ontology, natural language, and information systems: Implications of cross-linguistic studies of geographic terms*, in: M. Gould, R. Laurini, S. Coulondre (Eds.), *AGILE 2003: 6th AGILE Conference on Geographic Information Science*, Lyon, France, 2003.
- [119] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, A. Oltramari, *Ontologies library (final)*, *WonderWeb Deliverable D18*, ISTC-CNR, Padova, Italy, December 2003. URL: <http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf>.
- [120] C. Masolo, L. Vieu, *Atomicity vs. infinite divisibility of space*, in: C. Freksa, D. Mark (Eds.), *Spatial Information Theory*, *Proceedings of COSIT'99*, in: *LNCS*, vol. 1661, Springer-Verlag, Berlin, 1999.
- [121] P. Matsakis, J. Keller, L. Wendling, J. Marjamaa, O. Sjahputera, Linguistic description of relative positions in images, *IEEE Transactions on Systems, Man, and Cybernetics*, Part B 31 (4) (2001) 573–588.
- [122] N. Mavridis, D. Roy, Grounded situation models for robots: Where words and percepts meet, in: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2006.
- [123] G. Miller, P. Johnson-Laird, *Language and Perception*, Cambridge University Press, Cambridge, 1976.
- [124] R. Moratz, T. Tenbrink, Spatial reference in linguistic human–robot interaction: Iterative, empirically supported development of a model of projective relations, *Spatial Cognition and Computation* 6 (1) (2006) 63–107.
- [125] I. Niles, A. Pease, *Toward a standard upper ontology*, in: C. Welty, B. Smith, C. Welty (Eds.), *Formal Ontology in Information Systems (FOIS)*, Association for Computing Machinery, New York, NY, USA, 2001. URL: <http://projects.tekknowledge.com/HPKB/Publications/FOIS.pdf>.
- [126] I. Niles, A. Pease, *Linking lexicons and ontologies: Mapping WordNet to the suggested upper merged ontology*, in: *Proceedings of the IEEE International Conference on Information and Knowledge Engineering*, IEEE, 2003.
- [127] D. Oberle, A. Ankolekar, P. Hitzler, P. Cimiano, M. Sintek, M. Kiesel, B. Mougouef, S. Baumann, S. Vembu, M. Romanelli, P. Buitelaar, R. Engel, D. Sonntag, N. Reithinger, B. Loos, H.-P. Zorn, V. Micelli, R. Porzel, C. Schmidt, M. Weiten, F. Burkhardt, J. Zhou, *DOLCE ergo SUMO: On foundational and domain models in the SmartWeb Integrated Ontology (SWIntO)*, *Journal of Web Semantics* 5 (2007) 156–174.
- [128] P. Olivier, K.-P. Gapp (Eds.), *Representation and Processing of Spatial Expressions*, Lawrence Erlbaum Associates, Mahwah, NJ, 1998.
- [129] P. Olivier, J. Tsujii, A computational view of the cognitive semantics of spatial prepositions, in: *32nd Annual Meeting of the Association for Computational Linguistics*, New Mexico State University, Las Cruces, New Mexico, 1994.
- [130] A.G. Philpot, E. Hovy, P. Pantel, The omega ontology, in: *Proceedings of the ONTOLEX Ontologies and Lexical Resources Workshop at IJCNLP-05*, USC/Information Sciences Institute, Los Angeles, CA, 2005. URL: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.141.4699>.
- [131] I. Pratt, N. Francez, *Temporal prepositions and temporal generalized quantifiers*, *Linguistics and Philosophy* 24 (2) (2001) 187–222.
- [132] J. Pustejovsky, *The Generative Lexicon*, The MIT Press, Cambridge, MA, 1995.
- [133] J. Pustejovsky, *Lexical semantics and formal ontologies*, in: N. Guarino (Ed.), *Formal Ontology in Information Systems (FOIS)*, IOS Press, Amsterdam, 1998, pp. 328–336.
- [134] A. Rashid, B. Sharif, M. Egenhofer, D.M. Mark, 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) (1998) 215–246.
- [135] M. Raubal, M. Egenhofer, D. Pfoser, N. Tryfona, Structuring space with image schemata: Wayfinding in airports as a case study, in: S. Hirtle, A. Frank (Eds.), *Spatial Information Theory—A Theoretical Basis for GIS*, *International Conference COSIT '97*, Laurel Highlands, PA, in: *Lecture Notes in Computer Science*, vol. 1329, Springer-Verlag, Berlin, 1997, pp. 85–102.

- [136] G. Retz-Schmidt, Various views on spatial prepositions, *AI Magazine* 9 (2) (1988) 95–105.
- [137] R. Ross, H. Shi, T. Vierhuff, B. Krieg-Brückner, J. Bateman, Towards dialogue based shared control of navigating robots, in: C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, T. Barkowsky (Eds.), *Spatial Cognition IV: Reasoning, Action, Interaction*, International Conference Spatial Cognition 2004, Proceedings, Frauenchiemsee, Germany, October 2004, Springer, Berlin, Heidelberg, 2005.
- [138] R.J. Ross, Tiered models of spatial language interpretation, in: C. Freksa, N.S. Newcombe, P. Gärdenfors, S. Wölfl (Eds.), *Spatial Cognition VI: Learning, Reasoning and Talking about Space*, International Conference, Spatial Cognition 2008, Freiburg, Germany, in: *Lecture Notes in Artificial Intelligence*, vol. 5241, Springer, 2008.
- [139] R.J. Ross, C. Mandel, J. Bateman, S. Hui, U. Frese, Towards stratified spatial modeling for communication and navigation, in: *IROS Workshop From Sensors to Human Spatial Concepts 06*, Beijing, China, 2006.
- [140] U.-J. Rüetschi, S. Timpf, Modelling wayfinding in public transport network space and scene space, in: C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel, T. Barkowsky (Eds.), *Spatial Cognition IV: Reasoning, Action, Interaction*, International Conference Spatial Cognition 2004, Proceedings, Frauenchiemsee, Germany, October 2004, Springer, Berlin, Heidelberg, 2005.
- [141] SFB-360 Bielefeld Corpus, <http://www.sfb360.uni-bielefeld.de>, 2000.
- [142] H. Shi, T. Tenbrink, Telling Rolland where to go: HRI dialogues on route navigation, in: K. Coventry, T. Tenbrink, J. Bateman (Eds.), *Spatial Language and Dialogue*, Oxford University Press, Oxford, 2009, pp. 177–189.
- [143] M. Skubic, D. Perzanowski, S. Blisard, A. Schultz, W. Adams, M. Bugajska, D. Brock, Spatial language for human–robot dialogs, *IEEE Transactions on SMC, Part C* 34 (2) (2001) 154–167 (Special Issue on Human–Robot Interaction).
- [144] D.I. Slobin, Thinking for speaking, in: J. Aske, N. Beery, L. Michaelis, H. Filip (Eds.), *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society Meeting*, 1987.
- [145] D.I. Slobin, Two ways to travel: Verbs of motion in English and Spanish, in: M.S. Shibatani, S. Thompson (Eds.), *Grammatical Constructions: Their Form and Meaning*, Clarendon Press, Oxford, 1996, pp. 195–220.
- [146] B. Smith, Basic concepts of formal ontology, in: N. Guarino (Ed.), *Formal Ontology in Information Systems (FOIS)*, IOS Press, Amsterdam, 1998, pp. 19–28.
- [147] B. Smith, D.M. Mark, Do mountains exist? Towards an Ontology of Landforms, *Environment and Planning B (Planning and Design)* 30 (3) (2003) 411–427. URL: <http://ontology.buffalo.edu/smith/articles/Mountains.htm>.
- [148] S. Staab, R. Studer (Eds.), *Handbook on Ontologies*, Springer-Verlag, Heidelberg and Berlin, 2004.
- [149] M. Stone, Ontology and description in computational semantics, in: *Proceedings of the 3rd Workshop on Knowledge and Reasoning in Practical Dialogue Systems at the 18th IJCAI*, 2003. URL: <http://www.ida.liu.se/labs/nlplab/ijcai-ws-03/papers/stone.pdf>.
- [150] E. Stopp, K.-P. Gapp, G. Herzog, T. Laengle, T.C. Lueth, Utilizing spatial relations for natural language access to an autonomous mobile robot, in: B. Nebel, L. Dreschler-Fischer (Eds.), *Proceedings of KI-94: Advances in Artificial Intelligence*, 18th German Annual Conference on Artificial Intelligence, Springer-Verlag, Berlin and Heidelberg, 1994.
- [151] B. Swartout, R. Patil, K. Knight, T. Russ, Towards distributed use of large-scale ontologies, in: *Proceedings of the 10th Knowledge Acquisition for Knowledge-Based Systems Workshop*, Banff, Canada, 1996. URL: <http://ksi.cpsc.ucalgary.ca/KAW/KAW.html>.
- [152] L. Talmy, How language structures space, in: H. Pick, L. Acredolo (Eds.), *Spatial Orientation: Theory, Research, and Application*, Plenum Press, New York, 1983, pp. 225–282.
- [153] L. Talmy, Lexicalization patterns: Semantic structure in lexical forms, in: T. Shopen (Ed.), *Language Typology and Syntactic Description*, in: *Grammatical Categories and the Lexicon*, vol. III, Cambridge University Press, Cambridge, 1985, pp. 57–150.
- [154] L. Talmy, *Towards a Cognitive Semantics*, A Bradford Book, MIT Press, Cambridge, MA, 2000.
- [155] L. Talmy, The fundamental system of spatial schemas in language, in: B. Hampe (Ed.), *From Perception to Meaning: Image Schemas in Cognitive Linguistics*, Mouton de Gruyter, Berlin, 2006, pp. 37–47.
- [156] T. Tenbrink, Teaching an autonomous wheelchair where things are, in: K. Fischer (Ed.), *Proc. Workshop on 'How People Talk to Computers, Robots, and Other Artificial Communication Partners'*, Hansewissenschaftskolleg, Delmenhorst, April 21–23, 2006, SFB/TR8 Report 010-09-2006, Bremen, 2006, pp. 54–67.
- [157] T. Tenbrink, *Space, Time, and the Use of Language: An Investigation of Relationships*, Mouton de Gruyter, Berlin, 2007.
- [158] S. Timpf, Ontologies of wayfinding: A traveler's perspective, *Networks and Economics* 2 (1) (2002) 9–22.
- [159] A. Tyler, V. Evans, *The Semantics of English Prepositions: Spatial Scenes, Embodied Meaning and Cognition*, Cambridge University Press, Cambridge, 2003.
- [160] E. van der Zee, J. Slack (Eds.), *Representing Direction in Language and Space*, Explorations in Language and Space, Oxford University Press, Oxford, 2003.
- [161] C. Vandeloise, *Spatial Prepositions: A Case Study from French*, The University of Chicago Press, Chicago, 1991.
- [162] C. Vorweg, *Raumrelationen in Wahrnehmung und Sprache: Kategorisierungsprozesse bei der Benennung visueller Richtungsrelationen*, Deutscher Universitätsverlag, Wiesbaden, 2001.
- [163] S. Werner, B. Krieg-Brückner, T. Herrmann, Modelling navigational knowledge by route graphs, in: C. Freksa, W. Brauer, C. Habel, K. Wender (Eds.), *Spatial Cognition II—Integrating Abstract Theories, Empirical Studies, Formal Methods, and Practical Applications*, Springer, Berlin, 2000, pp. 295–316.
- [164] D. Wunderlich, How do prepositional phrases fit into compositional syntax and semantics? *Linguistics* 29 (1991) 591–621.
- [165] D. Wunderlich, M. Herweg, *Lokale und Direktionale*, in: *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, Walter de Gruyter, Berlin, 1991, pp. 758–785.
- [166] A. Yamada, T. Nishida, S. Doshita, Figuring out most plausible interpretation from spatial descriptions, in: *Proceedings of the 12th International Conference on Computational Linguistics (COLING-88)*, 1988.
- [167] H.D. Zimmer, H.R. Speiser, J. Baus, A. Krüger, Critical features for the selection of verbal descriptions for path relations, *Cognitive Processing* 1 (2001) 389–411.
- [168] J. Zwarts, Vectors as relative positions: A compositional semantics of modified PPs, *Journal of Semantics* 14 (1) (1997) 57–86.
- [169] J. Zwarts, Prepositional aspect and the algebra of paths, *Linguistics and Philosophy* 28 (6) (2005) 739–779.