The Role of Conceptual and Linguistic Ontologies in Discourse

John Bateman, Thora Tenbrink, Scott Farrar

October 1, 2006

Abstract

In this paper we argue that a clear division between two sources of information—one oriented to world knowledge, the other to linguistic semantics—offers a framework within which mechanisms for modelling the highly flexible relation between language and interpretation necessary for natural discourse can be specified and empirically validated. Moreover, importing techniques and results from formal ontology provides the formal underpinnings necessary for representing these sources of information appropriately. The result is a computationally specifiable model of dialogic interaction within which flexible discourse interpretation is seen as the result of inter-ontology mappings between our two sources of information. These mappings are dynamically negotiated according to the concrete discourse moves of interlocutors. The paper draws on our ongoing empirical studies in spatial discourse, where interlocutors jointly solve varieties of spatially-embedded tasks, and suggests that ontological formalization is of considerable benefit for constructing computational dialogue systems.

1 Introduction

It is now well accepted that language interpretation is inherently highly flexible and context-dependent. Linguistic terms and expressions typically need to be resolved against context in order to pinpoint their intended meanings. This process of flexible interpretation is often invisible to interlocutors; the unfolding dialogue and its context evidently provide substantial cues concerning just how underspecified meanings are to be filled out. Moreover, whenever there are difficulties in resolving interpretations, interlocutors are able to construct clarificatory interactions to home in and construct common shared interpretations with considerable precision. Several mechanisms have been proposed in the linguistic literature for accounting for this facility. However, these remain fragmented and oriented to specific cases. In this paper our concern will be to argue that a clear separation of two distinct levels of representation—one oriented to world knowledge, the other to linguistic semantics—can contribute significantly to providing generic mechanisms for specifying and exploring flexible dialogically-embedded discourse interpretation.

Our approach adopts results gained in formal ontological engineering and proposes that the two levels of representation required are each specified in terms of full-blown axiomatized formal ontologies. This provides a robust foundation for modelling decisions which can then be subject to finer empirical investigation. The formalization achieved also allows us to explore in more detail an approach in which discourse interpretation is captured by formal mappings between the two levels of representation. Fixing a particular discourse interpretation is then seen as equivalent to specifying particular inter-ontology mappings between a linguistically-oriented ontology and a conceptually-oriented ontology. Flexible discourse intepretation then requires that we set out just how particular inter-ontology mappings can be constructed on demand in the context of particular discourses. We will find many places where considerable further empirical research is required in order to fill out both the details of individual mappings and their construction and behaviour during unfolding discourses. Our claim will then be that having as a basis a model specified formally in the way we describe is a considerable aid to more focused empirical questions and hypotheses. In short, the model supports more focused investigations of the precise conditions holding on discourse interpretations.

The empirical basis of our work is drawn from the area of spatial discourse—i.e., discourse in which interlocutors act together in order to solve problems related to space, such as object and event localization, route descriptions, scene descriptions and the like. We are engaged in an ongoing research program which uses the inter-ontology mapping paradigm to pose questions and hypotheses that are then subjected to detailed empirical investigation. In the present paper we will draw extensively on these results. We will see that the area of spatial language in general benefits considerably from the modelling principles of ontological engineering. It is suggested that these benefits may also extend to many other commonplace areas of dialogic interaction.

We structure the paper as follows. First we highlight the issue of flexible interpretation—starting from some common examples and moving subsequently into the particular problems that arise in the spatial domain. On the basis of this, we argue the necessity of providing a clear separation between the two levels of representation we propose even though, traditionally, there has been considerable resistance to this separation. We then set out our approach to the two-level ontological modelling of space, drawing both from the extensive literature on spatial language and results coming from our empirical investigations. We show that particular interpretations correspond naturally to particular alignments between the two ontologies involved. This allows us to return to the observed flexibility of interpretation and characterize this as variability and changes in the applicable inter-ontology mappings. We conclude with an outlook for the future application of our paradigm for guiding empirical discourse research.

2 The inherent flexibility of spatial interpretation

In spatial communication and dialogic interaction concerning space, interpretative flexibility is particularly extensive. We will first focus on the nature of dialogues about space and identify some of the key tasks and strategies used together with the ensuing variability in the interpretation of spatial terms. We then discuss the various kinds of perspectives and reference systems that underly any interpretation of spatial terms independently of the current discourse task. Finally, we address the processes involved in identifying the intended message conveyed by a spatial term in dialogue.

2.1 Spatial language use and task dependence

Our investigation of spatial dialogue is based on empirical research aimed at establishing how interlocutors speak in general (Tenbrink, 2005a,b) and to spatially-aware assistance robots (e.g., Moratz and Tenbrink, 2006; Fischer, 2006) about spatial situations. In any dialogue concerning space, we can identify the following generic 'meta-tasks' that the speaker may be pursuing:

1. To describe configurations of objects, or 'scenes': describing a spatial scene is the task of establishing where objects are located with respect to one another.
2. To refer to an object by identifying its spatial relationship to one or more other objects.
3. To give route descriptions, i.e., enable the interlocutor to reach a certain goal location.

The first and third of these tasks have received extensive attention in the spatial language literature, the second less so. All of these tasks involve complex processes of establishing the correspondence between the verbal representation and the perceived world. Achieving these tasks within the context of dialogic interaction adds considerable opportunities for misunderstandings. Spatial dialogue is made particularly complex by the fact that human users employ diverse strategies for communicating about spatial relationships. The spatial expressions produced are themselves subject to a range of different interpretations and implications depending on the global task being pursued and the strategy for solving it adopted. To give a clear sense of this kind of variability, we set out for each of the three meta-tasks in turn the interpretation of the simple German spatial term: *links* ("left"). As we will describe below, our data is primarily taken from German dialogic interactions and so the examples presented are accordingly in German with English glosses.

First, in spatial scene descriptions, speakers commonly use **projective** terms like *vor/hinter* (in front/behind) and *links/rechts* (left/right) as in (1); these expressions locate objects in relation to one another with respect to a specific spatial axis.

(1) Die Tür ist links von dir.
'The door is to your left.'

The application of projective terms in spatial scene descriptions like these depends on the spatial relationship between the intended object, which we term the locatum, and a selected relatum. Unmodified projective terms may be used when the locatum lies directly along an identified axis (e.g., 'left'). With increasing angles, modifications of the spatial term and combinations of several terms become increasingly likely. Effects such as these have been described in terms of mechanisms such as 'spatial templates' (Carlson-Radvansky and Logan, 1997) and 'applicability fields' (Gapp, 1995; Zimmer et al., 1998). In addition to projective terms, spatial scenes can also be described using topological or path-related terms, and expressions denoting distance or in-between relations (cf. Tenbrink, 2005c).

Second, in tasks requiring referential identification, terms may be used rather differently. Referential identification requires that an object be singled out on the basis of contrast to other objects present (Herrmann and Deutsch, 1976). Such a contrast may be established both spatially and by other features of the objects involved. In some cases, the class type of the object alone may be sufficient if it is discriminative, or other discriminating features such as color, shape, and size may be required. Alternatively, if spatial projective terms are used for discrimination, they only need to be modified if other objects can also be described by the same spatial term. The largest area that can be assumed (and has been proposed) for applicability of a projective term (depending on the situation) is a half-plane (e.g., Herskovits, 1986, p181f.). Therefore, while the applicability areas of *left* and *front* are not mutually exclusive, those of *front* and *back* as well as *left* versus *right* are. It is probably for referential identification tasks that the option of using a full half-plane as a possible region of application for an unmodified projective term like *links* (left) is most relevant since more specific descriptions may not be necessary for successful communication (Vorwerg and Tenbrink, 2006).

Finally, in the route description task, speakers may refer directly to the goal itself as in (2a), using referential identification as just described and leaving path information unspecified. Alternatively, speakers may refer to the path or direction to be taken but not the goal entity, as in (2b). Full descriptions contain information about both direction and goal, as in (2c).[1]

(2) a. Gehen Sie zu dem Quadrat ganz hinten links.
'Go to the square all the way back and to the left.'

b. Gehe geradeaus, leicht nach links.
'Go straight ahead, slightly to the left.'

c. Gehen Sie geradeaus an den Quadraten links von Ihnen vorbei.
Es ist dann das hinterste Quadrat.
'Go straight ahead past the squares to the left of you. It is then the backmost square.'

Descriptions that specify the path or movement towards a goal often involve similar or even identical terms as those used in the other tasks; sparse utterances like "links" as a stand-alone term are not uncommon in on-line (i.e., accompanying) movement instructions. In this case, the term needs to be interpreted dynamically, i.e., it refers to a re-orientation towards the left side of the mover, or even a movement in that direction (Tenbrink et al., forthcoming). Deciding between these two options, and furthermore deciding about the quantity of the movement (turning angle as well as distance to be covered by motion), is not in all cases trivial. Again, these aspects depend on the current situation to a high degree.

Full descriptions may contain many kinds of additional information, such as detailed directional information, information about decision points (Denis, 1997), and reference to further entities such as sub-goals, obstacles, or landmarks. Often, route description tasks involve a structured environment such as a street network. In this case, an expression like "links" may commonly mean that the addressee should turn left at the next possible intersection, according to the actual options the street network offers. This includes the streets' angles: typically, the projective term *links* will not be modified or specified in more detail in a route instruction if there are no other competing streets on the left side of the current street (Klippel et al., forthcoming). The more complex the spatial situation, however, the more complex the

[1] These examples are slightly adapted from our corpus as described in Tenbrink (2005b).

reference to a turn will be (similar to the usage of projective terms in referential identification tasks); and as a result, the level of detail involved in interpreting the spatial term will be increasingly fine.

As we have seen from this brief overview on generalized situations in which an expression like "links" (left) may occur, this term (even as a stand-alone utterance) can be interpreted in any of the following ways:

- as a position directly on the prototypical left axis of a relatum (i.e., on a 90^o angle),
- a position on the left half-plane with respect to a relatum (i.e., within a range of 0^o to 180^o),
- anything in between these two extremes,
- a movement in the left direction of the mover,
- a re-orientation towards the left of the mover, or
- a turn into a direction specified by the environment - such as a street network - on the left side of the mover.

Thus, while the term 'left' may appear simple, when combined with other spatial descriptions and in the context of particular spatial tasks, even this term could lead to considerable confusion and error. Which of these interpretations is suitable in a given situation needs to be derived from various sources and this then forms one of the main problems that we address.

2.2 Perspectives and reference systems

One particularly significant source of variability in interpretation is due to the fact that all usages of projective terms involve commitments to both **perspective** and **reference system**. Perspective is defined as the particular point of view from which all relations are to be interpreted. While it is possible to encode the perspective directly in dialogue, as shown in bold in the gloss given in (3a), this strategy has rarely been encountered in our data. Instead speakers typically only refer explicitly to the relatum as in (3b), although this kind of information may also be omitted, as illustrated in (3c) (examples from our corpora).

(3) a. Gehe zu dem Objekt, welches sich von dir aus gesehen auf der rechten Seite befindet.
'Go to the object which **from your point of view** is on the right hand side.'

b. Gehe zu dem Kreis rechts von Dir.
'Go to the circle to **your** right.'

c. das Quadrat links
'the left square'

Reference systems can be succinctly characterized in terms of spatial *roles* that may be allocated to selected objects in the spatial scene. These roles include relatum, locatum, and origin. The origin is the position defining the perspective as just described. Three distinct types of reference system are commonly distinguished; these are termed, following Levinson (2003), **intrinsic**, **relative** and **absolute** reference systems. In the case of intrinsic reference systems, the relatum and the locatum both pick out the same object; this is the case in (3a) and (3b), which are also both suitable for describing the same spatial concept. In relative reference systems, however, the relatum differs from the origin. This is shown in the following example where the speaker serves as relatum while, as made explicit by the speaker, the adressee provides the perspective:

(4) Gehe zu dem Objekt, das, wenn Du mich anguckst, links von mir steht.
'Go to the object which, if you look at me, is located left of me.'

Absolute reference systems function differently in that they make use of specific linguistic expressions, such as **north** and **south**, that pick out directions in the environment independently of any selected perspective or relatum.

Although there is often in principle a relatively free choice between possible reference systems, there are differences in their typical situations of use. Intrinsic reference systems, for example, are generally employed when their relata can contribute asymmetries that allow for the allocation of distinguished parts, such as 'fronts' and 'backs', or 'left' and 'right'. Such asymmetries may also be functionally induced by means of the use to which particular parts of an object are put—such as the entrance to a building constituting its front—or be related to features considered relevant for perception—such as one's face defining a person's 'front'. In addition, objects in motion (or potentially in motion) can induce a further kind of perspective that is independent of intrinsic fronts or perceptual organs. Here, the perspective adopted can be described as though the moving object was viewed from the inside, so to speak, looking in the direction of motion. This permits even a completely symmetric object, such as a ball, to be ascribed front, back, right and left sides when in motion as shown in (5).[2] Note that (5) can also be interpreted with a *relative* reference system using an external viewpoint, so that the utterance is potentially ambiguous. It is not trivial to predict which of these options would be preferred in natural discourse.

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

The assignment of sides in the case of motion also holds for scenarios in which the moving entity itself serves as the origin of a relative reference system. In this case intrinsic features do not play a role, but rather the direction of motion, as in:

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

In our approach we consider any oriented object as capable of serving as the origin of both intrinsic and relative reference systems. That orientation can come about both by intrinsic features and by (potential) motion. Similar effects arise by functional ordering relations such as those induced by a queue, where sequentiality is induced by the interplay of temporal and spatial aspects (e.g., people waiting closer to the front end of a queue will be served earlier than people behind them).

[2]This and the following example are invented.

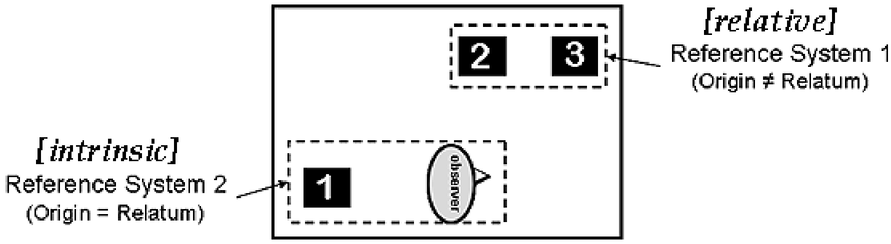

Figure 1: Conflicting Reference Systems: 'at the back / behind'

The basic problem of reference system ambiguity, or underspecification, is shown clearly in the following example. Consider the spatial situation depicted in Figure 1. To refer to one of the objects, the German adjective *hinter-* could be used as in (7):

(7) Geh zum hinteren Objekt.
'Go to the object behind.'

However, due to the availability of different reference systems, *hinter-* ('behind') is ambiguous and can be applied to at least two of the objects in the environment, namely, object 3 as well as object 1. It applies to object 3 if it is assumed that object 2 is used as the relatum of a relative (in this case, group-based, cf. Tenbrink and Moratz, 2003) reference system in which the observer is the origin. In this case, object 3 is situated behind object 2 in the sense that the view direction from the observer can (a) be shifted so as to lie along the row suggested by the positions of object 2 and object 3 and (b) encounter object 2 first. The real possibility of such an interpretation has been supported by our empirical work (Moratz and Tenbrink, 2006). Object 1, on the other hand, is situated behind the observer and, if the underlying reference system is intrinsic, can therefore also be referred to as *hinter-*, using the observer as both origin and relatum.

2.3 The problem: interpreting spatial reference in dialogue

The possibilities for interpretation illustrated briefly above show that one primary task during a spatial dialogue must be the negotiation of a reference system and the dynamic assignment of the various spatial roles necessary for understanding what is at issue. The main problem for both interlocutors and theories of spatial discourse interpretation lies in the fact that it is overwhelmingly the case that the linguistic surface forms employed do not directly reflect underlying reference systems.[3] Interpreting instances of spatial reference in natural discourse is therefore a complex task that involves both knowledge about general discourse principles, knowledge about the available options for interpretation, and the dynamic negotiation of suitable interpretations (such as underlying reference frames). That is, the speaker must impart his or her view of the world using language, and in such a way that the intended spatial relationships are recoverable based on shared contextual knowledge.

[3]There are, in fact, a number of additional conceptual options for establishing reference systems, described in more detail in Tenbrink (2005c).

However, although there is extensive variability in principle, that variability is not free-floating. Interlocutors commonly omit crucial details concerning which reference frame they are employing and the spatial role assignments adopted but their hearers do not react as if information has been withheld. Quite the contrary appears to apply: speakers and hearers continue interacting as if the problems alluded to in the previous sections were not there at all. When there is a difficulty, it is quickly corrected by means of clarification dialogues apparently without particular effort. Much of our empirical work is therefore aimed precisely at finding the particular sources of constraint that enable spatial dialogues to proceed so smoothly.

One source of clarification is the discourse history. For example, many speakers of German consistently use adjectives (as in 'die linke Schale', 'the left bowl') for group-based relative reference systems, whereas they employ adverbs and explicit relata for intrinsic reference systems (as in 'die Schale links von dir', 'the bowl left of you') (Tenbrink, 2005c). Thus, if the current discourse supports this default assumption, it can be assumed with relative safety that the ambiguity in (7) above is not an issue in this specific case: the intended object is almost certainly object 3 in Figure 1. In fact, our human-robot interaction data show that speakers are indeed rather consistent in their choice of reference systems, including their syntactic variants, at least in case of discoursal success (Moratz and Tenbrink, 2006). Thus, interpretation can be restricted throughout the discourse until misunderstanding or unresolvable ambiguity arises.

Our task now will be to provide a generic way of modelling discourse that provides a natural home both for this observed flexibility and for the mechanisms of negotiated constraint of that flexibility. This effort is therefore situated within current research that explores the general mechanisms at work in producing and interpreting natural language in diverse discourse situations. Particularly relevant to our concerns here are psycholinguistic approaches concerning knowledge about speakers' preferences in object reference (Herrmann and Deutsch, 1976) and the theory of *interactive alignment* (Pickering and Garrod, 2004), whereby the space of possibilities theoretically available in any dialogue is constrained by the fact that speakers come to adopt forms and interpretations that have been established earlier in the interaction. This means that once a form-meaning mapping has been established, interlocutors will 'by and large' stay with this mapping until there is sufficient reason to change (cf. Tversky et al., 1999). These approaches leave many questions unanswered concerning just how the utterances of speakers can lead to the level of sustained detail required for successful spatial interaction. Our own approach, therefore, will be to specify the linguistic and conceptual possibilities involved in spatial language in sufficient detail as to make possible explicit mapping mechanisms between the levels of description required. And to do this, we must first set out the formal ontological frameworks that we apply.

3 Ontological modelling for dialogue

In their early work demonstrating the flexibility of language interpretation in dialogue, Garrod and Sanford considered shared mental models constructed by interactants in their discourse to provide a way of mediating between language and the spatial world that they inhabit. Such models were described thus:

> "Basically, a mental model of a situation consists of two things: (1) some set of autonomous objects which map onto the situation and give it an 'ontology'... and (2) a tight knit set of relations between the objects in the domain which capture the 'topology' of the situation..." (Garrod and Sanford, 1988)

In the work we describe in this paper, we have taken this mention of the construction of an 'ontology' particularly seriously. We apply significant results that have been achieved in the design principles of formal ontology explicitly to the construction of an account of dialogue interpretation—that is, 'ontology' is taken in a precise technical sense rather than as a term for 'some structuring of the world'. This approach calls for fine-grained modelling that, we suggest, can lead to a solid basis for further empirical investigation. The issues we address in this section therefore revolve around ontologies and their use. We begin with a brief definition of ontology and our motivation for considering at least two distinct 'levels' or 'strata' of ontology for an account of discourse. We then introduce the particular ontologies that we are using and developing further and the position that space takes up within these. This provides the formal background to our consideration of actual spatial language usage as revealed by our empirical investigations.

3.1 What is an ontology and why do we want (at least) one?

For modelling systems in which rich background knowledge is essential, there is now an increasing need for formal, computationally viable descriptions of the domains of operation of those systems. This requirement shares tantalising similarities with the traditional goals of philosophical Ontology. Just as Ontology investigates and characterises the necessary structures of reality—the world in which humans act and operate—so computational ontology needed to characterise the structure and properties of the domains in which computational constructs such as intelligent systems were to operate. Computational ontology thus raised the issue of characterising these common-sense features of the world that we take for granted. Early computational ontologists, such as Patrick Hayes (cf., e.g., Hayes, 1979) began setting out formal systems intended to capture the 'naive physics' of the everyday world.

The particular approaches that we discuss here have grown out of the extensive experiences gained since then in this area. In particular, ontology designers have come to concern themselves increasingly with **foundational ontologies**—ontologies that are sufficiently general that they describe the essential characteristics of objects, properties and relations as such and in *all possible domains*. Moreover, formal methods are being developed in order to relate such foundational ontologies to particular domains and to deconstruct ontologies into modular components. Both of these areas are central to the account we present here.

The development of ontologies in fact has several points of contact with the tasks of discourse modelling. Discourse modelling is known to be a knowledge-intensive process; many accounts of discourse interpretation rely crucially on extensive world-knowledge being available. Ontological approaches set out to make that knowledge as well-founded and organized as possible and so serve to provide strong constraints within which any particular modelling needs to operate. Formal ontologies also provide sufficient content to allow sophisticated reasoning concerning the domains represented. This is therefore also of use for theories of discourse in that interlocutors quite clearly build on a rich understanding of how the entities in their worlds interrelate and are interdependent. This application of ontologies for discourse processing is therefore finding growing attention and several dialogue architectures have now been proposed drawing on ontological principles (e.g., Beun et al., 2004).

3.2 The need for distinct levels of ontology

Once the decision has been made to employ ontological components in an account of dialogue, there are still several alternative paths that can be taken. One of the most significant for the subsequent design of a discourse processing model is the question of how much work a particular ontological organization is to do within the account. Some systems designed for dialogue processing take the position that since an ontology provides for interoperability between components, adopting a single ontology for all aspects of its knowledge organization is a beneficial architectural decision. A similar position is reached by default, when the working assumption is that there is only one 'kind' of knowledge to be maintained and this is where diverse sources of information are combined.

Whereas the use of ontological principles for organizing domain is highly beneficial because of increased rigor and consistency it brings for the modelling, regardless of the precise approach taken, assuming such an ontology takes us back to the problem of the flexibility of language. Adopting an ontology with well-founded design principles has the side-effect of showing just how flexible the usage of language is. In order to deal with this flexibility, two traditions have emerged in the application of ontologies for discourse processing and the construction of computational systems for discourse processing built on them. The first adopts only one level of ontological organization and takes on the task of relating linguistic behaviour, no matter how varied, to this level of representation. In contrast to this, the second assumes that ontological representations need to be layered, or stratified, in order to capture very different kinds of organization with differing motivations.

Although more details of the arguments for 'layering' our ontologies in this way are given in Bateman (to appear), we can readily see something of the need for this layering in the following observation by Hobbs concerning an everyday spatial object, 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." (Hobbs, 1995, p820)

When we consider how language treats the 'road' at issue, we see again a high degree of variability. Such varying linguistic expressions as 'on the road', 'in the road', 'at the road', 'along Route 66', etc. are all possible. Problematic is the fact that *linguistically* the usage of these diverse prepositions appears to commit to entities that behave as if they have distinct dimensionalities—prepositional phrases with 'in', for example, strongly suggest three-dimensional containment.[4] We need then to relate an entity in the world, an actual road, with very different appearing versions of that entity that might appear in discourse. Ontologically, however, it is not appropriate that an entity, such as a road, change its ontological

[4]This is quite a reliable construal in English; other languages, even quite close languages such as German, require differing treatments.

characterization (container, surface, point, etc.) according to the way it is spoken about. A basic category such as dimensionality should not behave in such a non-rigid fashion.

Whereas approaches in the first tradition above must take on the task of relating form and referent by establishing more or less complex mechanisms for selecting a form that commits to a dimensionality that differs from that of the referent, the second tradition can approach the problem in a different way. The adoption of a further layer of ontological description, that we term a **linguistic ontology**, offers the possibility of organizing an additional layer of informational structuring that respects linguistic categorizations rather than world categorizations. We will see below that this is precisely what is required in order to cope with the complex interplay of linguistic and non-linguistic commitments within spatial discourse.

Allowing both a conceptual and a linguistic ontology fixes both ends of a mapping between ontological organizations: instead of a looser collection of linguistically determined elements, such as the partial tree structures with composed lexical semantics found in lexicalized grammar approaches, which need, somehow, to be brought into relation with a domain ontology, we have two clearly defined ontological layers, each with their own clearly defined design principles. The mapping issue then can be restated in terms of inter-ontology relationships. And the flexibility of the mapping is reconstructed in terms of dynamic inter-ontology mediation. The primary benefits of this approach, which we will show in detail below, are two-fold: first, we achieve a highly modular approach to linguistic modelling, interpretation and production, in which full justice can be given to the complexity of the issues involved; second, we can import results from the area of ontological engineering, an area now undergoing dramatic development as the need for ever more ontologies grows.

3.3 The ontologies we adopt

We have suggested that we need two kinds of ontological organization for dealing with the flexibility shown by natural language use: conceptual ontologies, that are constructed so that their categories are language-neutral, motivated according to cognitive or other non-linguistic criteria, and linguistic ontologies, which are responsive to the classification and relationships constructed by language as a way of construing the world.

In Farrar and Bateman (2004), we provide an extensive review of the current state of the art in conceptual ontologies and propose the adoption of the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE: Masolo et al., 2003), developed originally as a part of the WonderWeb project,[5] as a solid foundation for further work adopting ontological methods. DOLCE is an explicit attempt to take as much as possible from the results in *formal* ontology from the last 10 years and to build a comprehensive upper ontology respecting those requirements. The purpose of the ontology is to support a wide range of cognitive and linguistic modelling and so is particularly appropriate for considering issues of discourse processing also.

Following from its grounding in formal ontology and aim at providing a *foundational* ontology, DOLCE is most closely concerned with sorting out the upper few levels in the ontology in a manner that is philosophically and logically defensible. As its designers state, a positive goal

[5] IST Project 2001-33052 WonderWeb: Ontology Infrastructure for the Semantic Web

of the ontology is to be 'minimal'—i.e., it should be "as general as possible, including only the most reusable and widely applicable upper-level categories." (Masolo et al., 2002, p7). The most fundamental division made in DOLCE is accordingly that between entities that unfold in time, called perdurants (loosely: events), and entities which are present 'all-at-once' in time, called endurants (loosely: objects). Both such entities can be the bearers of qualities. DOLCE distinguishes between the particular qualities that inhere uniquely in particular entities and more abstract quality 'spaces' that define how particular qualities may be related to one another. Thus, a particular colour may inhere in a physical object, but this colour is itself drawn from the abstract region of a 'colour space' that defines the ontological properties of colours as such and allows distinct colours to be compared. DOLCE takes a distinctive view of spatial locations by placing 'space' alongside abstract regions such as 'colour'. Thus, just as a physical object may have a quality of colour situated within a colour quality region, it may also have a quality of 'being-located' situated within a spatial quality region. We will see below that this has allowed us to define a highly modular and flexible approach to formalizations of space that also support flexible discourse processing.

Linguistically motivated ontologies contrast sharply with conceptual ontologies in that their categories are motivated strictly according to the lexicogrammar and semantics of natural languages. The first ontology of this kind was constructed for a computational system for automatically generating natural language texts and was called the **Penman Upper Model** (Mann et al., 1985). The purpose of this ontology was to serve as mediator between, on the one side, domain knowledge that was to be expressed in natural language and, on the other, linguistic semantic knowledge capturing regularities of linguistic expression. Thus, categories were accepted within the Upper Model only when they were also motivated by distinctive patterns of linguistic realization. Knowing which upper model category had been selected then provides a strong set of constraints concerning the language that may occur and so could significantly aid the generation process. The Upper Model subsequently went through several phases of development resulting in the **Generalized Upper Model** (GUM: Bateman et al., 1995), the most extensive linguistically motivated ontology of its kind. This is the ontology that we have adopted for our linguistic ontology and, as we will show below, we are currently augmenting it further with a rich collection of spatial concepts in order for application in robust spoken dialogue systems involving spatial communication.

Both of the adopted ontologies are formally represented. DOLCE is axiomatized in first-order logic with modality; GUM is expressed as a backbone in description logic (OWL-DL) with additional annotations in first-order logic. We will see the use of these annotations for spatial representation below.

3.4 Ontological modularity and spatial representations

In this section, we set out the position of space and spatial relationships within our two adopted levels of ontology.

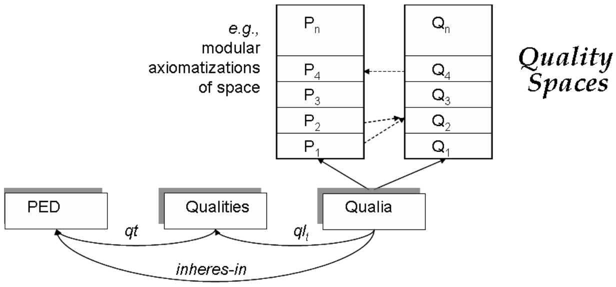

Figure 2: Definition of modular spatial accounts building on the foundation provided by the DOLCE ontology

3.4.1 Our treatment of space within DOLCE

DOLCE does not itself impose a strong view of how space should be modelled. The axiomatization binds together physical objects (endurants) and spatial qualities but, beyond presupposing a general mereological organization for any spatial model adopted, does not say how the spatial qualities are to be related to one another. In Bateman and Farrar (2004), we proposed a 'plug-and-play' approach to spatial ontologies whereby distinct kinds of formalizations of space could be adopted as required for particular tasks. This view is taken further and given a more integrated logic-based formalization in Bateman et al. (submitted); we depict this modular approach to spatial ontology graphically in Figure 2. Formalizations of a variety of qualitative spatial calculi have now been presented in, for example, Wölfl and Mossakowski (2005), and each of these can be incorporated as modules for the spatial ontological components shown in the figure.

The use of a powerful logic-based formalization of our ontological modules supports an interpretation of relations between our components as *theory morphisms* (cf. Bateman et al., submitted), allowing us to state precisely how configurations expressed in distinct components may be related. This is similar to the approach developed for relating distinct linguistically motivated ontologies in terms of *back-and-forth mappings* by Blackburn et al. (1996). In general, a theory morphism holds when the theorems following from the axioms of one theory can also be proved from the axioms of another theory under a specified translation between the constants, categories and predicates of the related theories. We will illustrate this below when we 'export' spatial information organized in terms of the linguistic ontology into corresponding configurations organized in terms of the conceptual ontology. This export/import relationship defines a mapping between ontologies and also constitutes a discourse interpretation of the linguistic configurations concerned. The precise formalizations of the spatial calculi adopted are, however, for our present purposes immaterial and we will present the mechanisms involved in more informal terms. In addition, we will focus the discussion particularly on issues relevant to spatial discourse processing by setting out how the well-developed view of the semantics of some aspects of natural spatial language presented by Eschenbach (1999) can be seen as a logical forerunner of the multi-ontology approach to spatial language that we present here and provides a good starting point for unravelling some of the flexibility shown above.

In Section 2.2, for example, we discussed the problem of the systematic ambiguity involved in the application of projective terms such as *left/right*, *in front of/behind*, etc. As shown there, each occasion of a projective term typically allows for the interpretation in relation to more than one reference system. Additionally, each projective term needs to be interpreted with respect to a region of application, which (as discussed in Section 2.1), may be anything from a line to a half-plane. To address these issues, Eschenbach (1999) separates out two areas of description necessary for the interpretation of these terms. She distinguishes between the direct contribution of the common or "core" linguistic semantics of the terms, which is constant across contexts, as opposed to the concepts that are required to interpret these terms in a concrete context (Eschenbach, 1999, p339). In her framework, an *abstract* geometric structure is imposed on the *concrete* spatial situation, depending on the underlying reference system (intrinsic, relative, or absolute), in order to yield a definite interpretation.[6]

Concerning regions of application, in Eschenbach's approach, *in front of* and *behind* are treated in terms of half-lines, whereas *right* and *left* are treated in terms of half-planes. Thus, the use of the lexeme *front* establishes a particular geometric structure in which there is an *orientation* that establishes a direction along which something can lie. Entities lying on the orientation thereby are assigned an ordering. The spatial expression itself then positions the spatial objects F (figure) and G (ground) on the orientation provided by the frame of reference. Crucial for understanding what is involved here is the understanding that the geometric structure provides precisely the abstractions that are necessary for assigning a semantics to the expression *front* and no more.

On the other hand, *left/right* (according to Eschenbach) require a geometric structure in which a half-plane is identified with reference to one of the entities related and the other is positioned within that half-plane. Clearly, the abstract constructions of an ordered orientation, on the one hand, and half-planes and containment, on the other, deal with very different geometric features. Figure 3 illustrates the account so far by showing how a single configuration of spatial objects (shown in the center) can be assigned an abstract geometric structure in these two different ways, depending on the reference system that is intended in the verbalization of the situation. On the lefthand side of the figure we see the assignment with respect to *in front of/behind* in terms of a half-line, and on the right, *left/right*, in terms of a half-plane. According to our discussions above, both half-lines and half-planes are possible as extreme variants of interpretation for any projective term, depending on the situation; ordering, however, as such seems only to be possible with the frontal axis.

Eschenbach is then able to specify just where the ambiguities arise, namely, in the assignment of roles within a reference system. For example, as explained in Section 2.2 the German lexeme *vor* (*in front of*) can be used in an intrinsic reference frame in which the relatum's intrinsic sides are used as basis for the establishment of a spatial direction. Alternatively, in a relative reference system the observer's line of vision is used as an origin for reference. Thirdly, a motion trajectory may impose a possible reference system on the objects. Eschenbach's proposal for these three options looks as follows (Eschenbach, 1999, p341).

(8) a. $vor_i \leftrightarrow\ < (o, Loc(F), Loc(G)) \wedge \mathit{Front\text{-}back\text{-}axis}(o, G)$

[6] Eschenbach introduces a systematic distinction between the terms *spatial reference system* and *frame of reference*, which we do not adopt here to avoid confusion since these terms are traditionally used interchangeably.

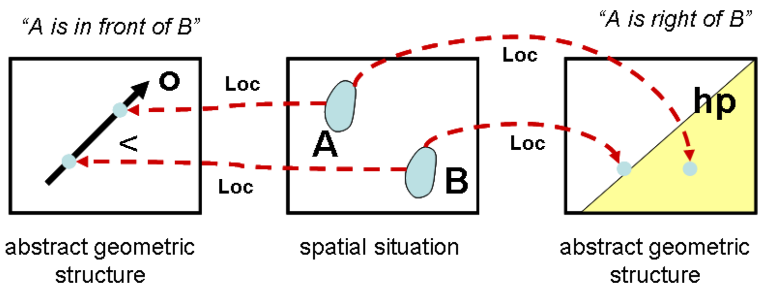

Figure 3: Two alternative ways of interpreting spatial expressions according to Eschenbach (1999)

b. $vor_r \leftrightarrow\ < (o, Loc(F), Loc(G)) \wedge \textit{Line-of-vision}(o, V)$

c. $vor_m \leftrightarrow\ < (o, Loc(F), Loc(G)) \wedge \textit{Trajectory}(o, X, G)$

In each case ordering is with respect to the identified orientation/direction o as given in Figure 3. The only difference is in how that orientation is fixed: according to an *intrinsic* axis of the ground (case 8a), a line of visibility defining a *relative* positioning (case 8b), or *movement* along some trajectory (case 8c). Eschenbach states: "The context provides the reference system (o) that yields the oriented line as additional parameter" (Eschenbach, 1999, p341).

As research since Eschenbach's publication has shown, the variability in the assignment of roles (or parameters) in spatial reference systems is even greater than previously assumed (Tenbrink, 2005c). There are also issues concerned with whether the distinction shown in Figure 3 actually corresponds to a difference between 'left/right' and 'front/back' or to different *uses* of the two sets of terms. Eschenbach's proposal is that the frontal axis is interpreted via an oriented axis, while the lateral axes induce half planes. In fact, analysis of our data suggests instead that there is little empirical foundation to the claim that the semantics and applicability of the terms are fundamentally different. That is, there is no proof for a fundamental distinction between linearity with the frontal and planarity with the lateral axis; all the projective terms are interpreted under the influence of discourse context, spatial templates, and all the other factors that we have seen above.

Nevertheless, we see the approach taken by Eschenbach as creating enough theoretical 'space' in order to describe both ambiguities and constants in the semantics of projective terms. Moreover, we can readily relate it to our ontological modules by considering any axiomatization of the properties of, in Eschenbach's terms, abstract 'geometric' space as a potential module of our ontology of space. The two kinds of information proposed by Eschenbach to distinguish what she terms the 'concrete' from the 'abstract' levels of space can then be straightforwardly related to the account of space that we are employing from the DOLCE ontology introduced above as follows.

The concrete spatial objects can be the physical objects (physical endurants) themselves, i.e., collections of physical objects with their necessary and ontological dependent qualities (such as location). The abstract geometric structures correspond almost exactly to the DOLCE abstract space region that allows locations to be labelled and compared. Formally, we can state the following mappings (cf. Figure 2 above):

Concrete spatial situation	$\mapsto$	$\{$x : DOLCE:*PED*$\}$
Abstract geometric structure	$\mapsto$	DOLCE:*Quality Space*
Loc	$\mapsto$	DOLCE:ql_t

This connection allows us to import Eschenbach's treatment directly into the generic foundational ontology adopted in our work. It also allows us to proceed to the most significant question for our current purposes: i.e., just how is the particular reference system and the corresponding abstract geometric structure appropriate for some usage identified?

3.4.2 Our treatment of space within GUM

We have extended the earlier versions of the Generalized Upper Model in the area of space and spatial relationships by focusing particularly on how information concerning space is expressed in both German and English. The details of this development are set out in depth in Bateman et al. (2006). Most significant here is the fact that the particular categories developed and their inter-dependencies are responsible both 'downward' to their linguistic realizations and 'upwards' to abstract spatial characterizations. The downward direction of responsibility has always been present in the design of the various versions of the Upper Model since, as we saw above, this constitutes one of its main motivations. The upward direction is, however, new in that it brings to bear the requirements of formal axiomatization imported from our commitment to principles of ontological engineering. The consequence of this double source of constraint is that we can relate particular uses of linguistic forms and constructions to particular abstract spatial representations. Each such representation makes a *minimal* commitment to the characterization of space that the linguistic term requires. As we shall see, these abstract characterizations are very similar to Eschenbach's abstract geometric spaces that we set out in diagrammatic form in Figure 3 above.

The spatial component that we have developed for the Generalized Upper Model is shown in Figure 4. Classes are defined in order to capture observations that certain groups of linguistic constructions behave in similar ways. For example, the spatial relation 'left', when used in examples such as (1) above, is a realization of the linguistic ontology category LeftProjectionExt(ernal). This is a subcategory of LateralProjectionExt (thereby grouping together 'left' and 'right'), which is itself a subcategory of HorizontalProjectionExt (grouping together 'left/right' and 'front/back'). Moving further to the left in the diagram, we see that these categories are themselves classified under ProjectionRelation and Disjointness. Each of these categories is related to a specified set of lexicogrammatical constraints concerning possible linguistic expression as well as making semantic commitments for their abstract spatial interpretation. This takes a characterization of the form given by Eschenbach and extends it to provide an additional criterion for organizing the concept/relation hierarchy of the linguistic ontology.

We can illustrate the diverse abstract geometric commitments of the GUM categories graphically as set out in Figure 5; each diagram in the figure should be seen as a graphical summary of a modular component of a complete axiomatization. We shall refer to these abstract geometric spaces in our concrete interpretation of dialogue extracts taken from our data below.

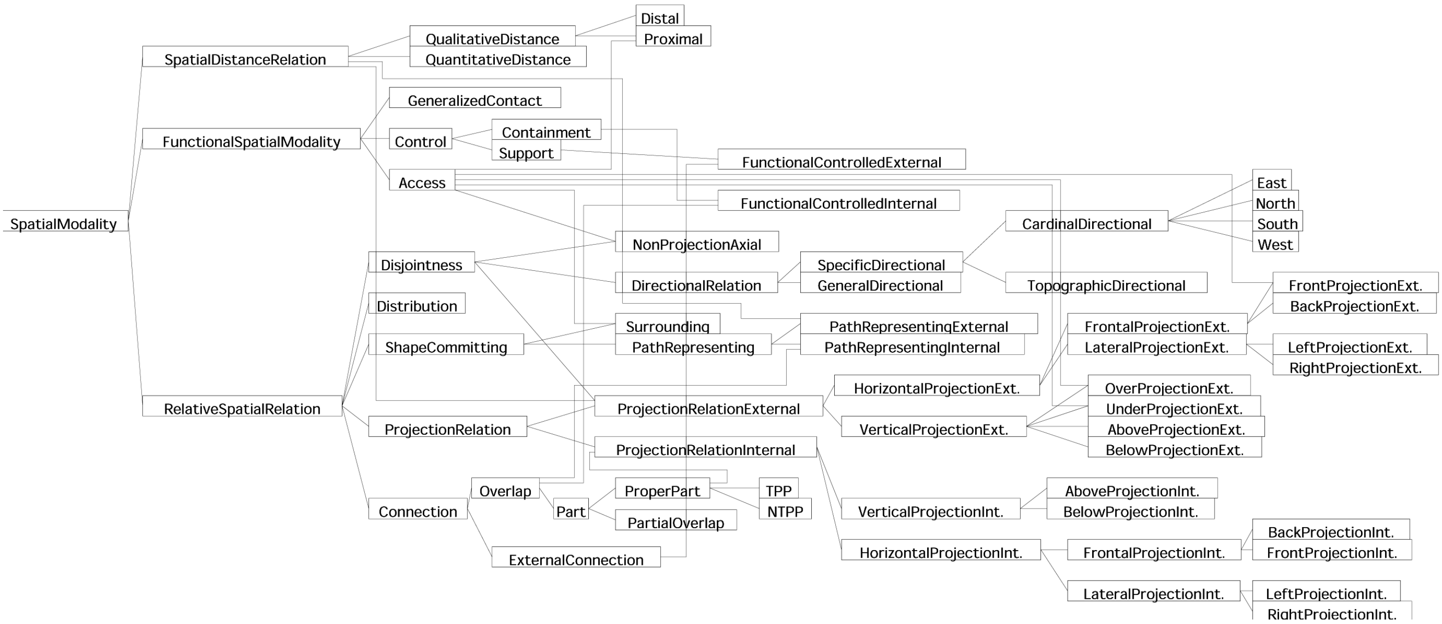

Figure 4: Spatially relevant extract from the Generalized Upper Model

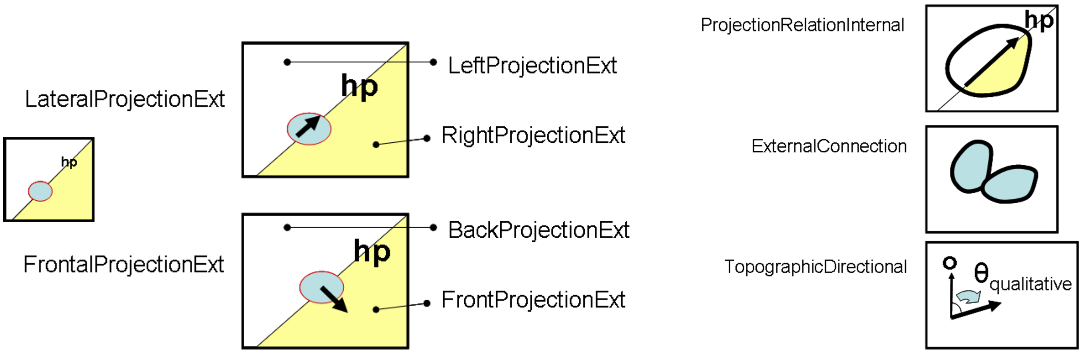

Figure 5: Correspondences between GUM spatial categories and abstract geometries

The diagrams on the left of the figure are directly related to Eschenbach's concerning half-planes shown in Figure 3 above. We now add to this, however, the *orientation* of the relatum, shown as a small arrow within the entity lying on the axis separating the half-planes. This captures the fact mentioned above that we consider left/right and front/back to share the basic property of decomposing the environment into half-planes and do not consider them as creating fundamentally different abstract geometries.

The diagrams on the right of the figure illustrate some further categories. One configuration common in our data is the ProjectionRelationInternal class which covers a similar range of lateral and frontal options but with respect to a specified bounded region: this allows expressions involving spatial 'parts' such as 'the left of the room', etc. How the region is allocated an orientation is not the concern of the abstract geometry, but is presupposed by its application. Finally, ExternalConnection corresponds to the straightforward connection relation of region-based spatial calculi and is often expressed by a preposition such as 'at', while

TopographicDirectional allows the specification of directions by means of qualitative 'angular distances' with respect to an arbitrary origin.

Each diagram sets out requirements for an embedding into the actual spatial context in terms of the origins, relata and orientations that need to be established for them to be applicable. The use of GUM during discourse processing is therefore as a mediator between concrete linguistic forms and semantic classes which invoke particular abstract geometric models. The task of contextualization is then to relate these abstract geometric models to particular spatial quality regions within the conceptual ontology and its characterization of the spatial environment. We will now show that *dialogic interaction* plays a very particular role in fixing the selected reference systems and the alignment between spatial entities and their abstract counterparts in geometric structure.

4 Spatial dialogue interpretation within the two-level ontology paradigm

In this section, we bring the discussions above, the two ontologies adopted, and the formal mechanisms for inter-relating ontologies to bear on actual examples from our corpus of spatial dialogues. The focus first will be to show that the variability that we have illustrated above is a regular and natural phenomenon in our tasks also and to set out how individual anchorings of linguistic and conceptual ontology decisions can be characterized. We then show how the account lets us move flexibly between mappings that speakers have adopted in order to cover the cases of adaptation and alignment of spatial interpretations within spatial discourse.

The particular data that we draw on in this section is taken from one series of experiments involving human-robot interaction. In this experimental setup, we confront our participants with a specifically designed spatial situation and ask them to perform a certain spatial task in interaction with a robot they do not know—i.e., they do not know exactly which functionalities can be expected. Thus, we create a controlled situation for linguistic data elicitation without controlling or limiting directly the range of linguistic choices that the speakers make (cf. Fischer, 2003).

Whereas in this paper we use the data obtained to show how inter-ontology mappings can characterize the situated interpretation of the language occurrences, we are also applying the data directly to the construction of working human-robot interaction systems. This then feeds back into a cycle of model refinement, since the linguistic and spatial behaviour of our robotic agents often do not match their users' expectations. This is a valuable way, therefore, of evaluating to what extent our interpretations correspond to those produced by human interactants.

4.1 Room descriptions

In the situation addressed here, our participants were working with an 'autonomous wheelchair', called Rolland (Lankenau et al., 1998). They had been told that the wheelchair was equipped with natural language processing capability, and their task was to describe to the wheelchair the layout of objects in a room. They were also told that once they had done this, Rolland

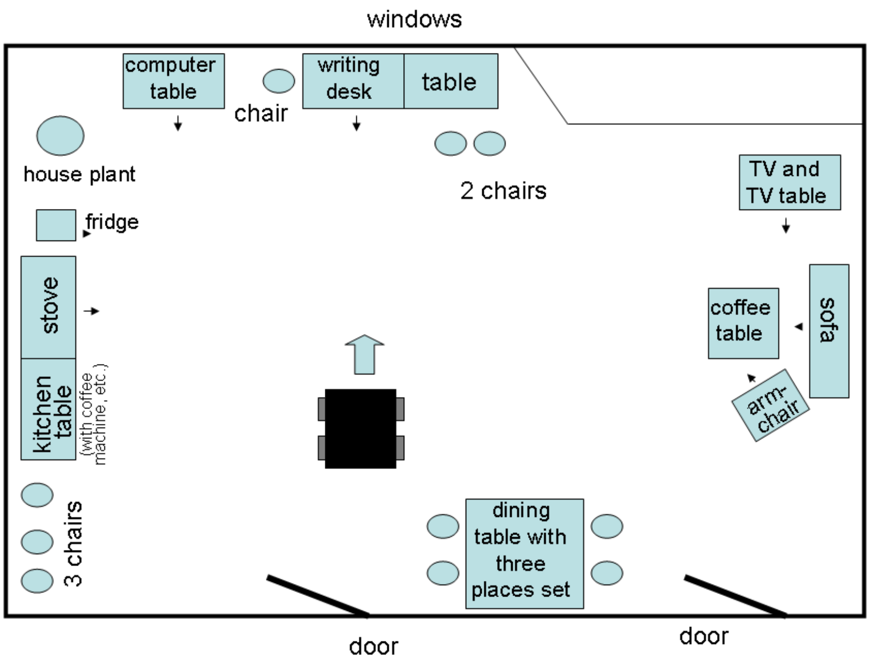

Figure 6: Room layout for the room description task

would be able to respond to voice commands for moving about the room. In this scenario we did not employ a real natural language processing system, but stored the participants' utterances for transcription and analysis. The layout of the room for all of the discourse extracts discussed in this section was as shown in Figure 6.

In one phase of the experiment, the participants had to give a running description of the position of the objects in the room to sum up what the robot wheelchair should have learnt. The wheelchair did not interact during this description (but see Tenbrink et al., 2006). The wheelchair and the participant seated on the wheelchair were always positioned as indicated in the figure, slightly left of center and down, and did not move during their description. Our first discourse extract in this scenario is R020. The description is divided into informational chunks for ease of reference. We present an approximate English gloss in the righthand column.[7]

[7]The transcription shows pauses: -, –, and (n) in seconds as well as filled hesitations, corrections and re-starts (+). Approximate intonation contours in terms of final falls and rises is shown with question marks and full-stops/periods.

Dialogue: R020 German

1	äh, - ja. - ähm, also von hier aus genau gegenüber ist mein Schreibtisch.	*er, so from here exactly opposite is my desk.*
2	und daneben links daneben ist mein Computer 'n Meter daneben vielleicht links daneben.	*and next to that left of that is my computer one meter apart*
3	(breathing) ähm, (1)	*er*
4	daneben an der Wand, is' meine Küche? - so s+ kommt zuerst mein Kühlschrank ganz rechts, –	*next to that at the wall is my kitchen, first there is my fridge all the way to the right*
5	so in der Ecke wo auch meine Blumen ge+, Blumen, meine Gießkanne, so sacht man, meine Gießkanne steht, –	*in the corner where there's also my flowers, my watering can*
6	und da is' mein Herd, und - der Tisch mit meiner Zuckerdose, ähm is' äh, auch auf der Seite von Kühlschrank? -	*and there is my stove, and the table with my sugar bowl is also at the side of the fridge*
7	links daneben,	*left of that*
8	(breathing) ähm, -	*er*
9	ja, - und äh, - also sozusagen hinter mir? aber links, äh, is' die Eingangstür? - klingelt, (1)	*yes, and er - so to say behind me but left, er, is the entrance door*
10	(breathing) dann äh weiter rechts neben mir gleich, (1) is' mein Esstisch? (1)	*then further to the right of me directly is my dining table*
11	und ähm, – so, - na wie is' es, von mir aus so sagen wir mal , – so zwischen äh zwölf Uhr und drei Uhr so is' genau der Fernseher . –	*and er, how is it, from my position let's say between er twelve o'clock and three o'clock there is directly the TV set*
12	(breathing) und ähm, rechts daneben äh steht mein Sofa, - da kann ich mich ja nich' drauf setzen, - nur die Gäste, -	*and er, to the right of that is my sofa, which I cannot sit on you know, only the guests*
13	(breathing) und auch mein äh mein, – schöner, Couchtisch, –	*and also my, er, beautiful coffee table*
14	(breathing) ja, - äh, -	*yes er*
15	und in der Mitte is' viel Platz, so d+ also steht alles schön an den Wänden so dass ich in der Mitte rumfahrn kann.	*and in the middle there is much space, so everything is placed nicely at the walls so that I can drive around in the middle.*

We will focus here on the spatial component of this description by picking out just the semantic analysis according to the categories of the Generalized Upper Model. The spatial relations of GUM are used in configurations that include specifications of relatum, locatum and adopted spatial relation (i.e., a category taken from the portion of GUM shown in Figure 4). This then includes the assignment of spatial roles as described in Section 2.2. The analysis is shown in Table 1. In terms of our linguistic ontology, therefore, the linguistic spatial terms that occur are organized into distinct classes and it is these classes that record the information as to precisely what kinds of reference frames and perspectives are entailed (cf. Figure 5 above).

As we can see, the structure of the description is relatively straightforward and progresses around the room loosely counter-clockwise (see Figure 6). What is consistently the case, however, is that we very rarely receive any indication of the perspective and type of reference frame that is being used. This is as described above, where we have generally found that it is precisely this rather critical information (as far as an umambiguous interpretation is

utterance	*locatum*	*relatum*	*GUM-category*
1	desk	self	NonprojectionAxis: opposite
2	computer	desk	LeftProjectionExt [distance: 1m]
4	kitchen	computer	HorizontalProjectionExt: next
4	kitchen	wall	ExternalConnection: at
4	fridge	kitchen	RightProjectionInt: rightmost
4	fridge	corner	Containment: in
5	houseplant	corner	Containment: in
6	stove	'there'	ExternalConnection: at
6	{stove, kitchen table}	fridge	HorizontalProjectionExt: side of
6-7	{stove, kitchen table}	fridge	LeftProjectionExt
9	entrance	self	BackProjectionExt
10	dining table	self	RightProjectionExt
11	TV	self	TopographicDirectional: 12-3 o'clock
12	sofa	TV	RightProjectionExt
13	coffee table	TV	RightProjectionExt

Table 1: Linguistic ontology characterization of the spatial content of dialogue R020

concerned) that is *usually* ommitted in natural descriptions. This strongly suggests that speakers do not generally see this information as missing at all. In addition, not even the relatum is always given explicitly, as in the case of 'ganz rechts' (interpreted as the relation 'rightmost'), where the conceptualized relatum is probably the kitchen area as a whole. This can be inferred from the discourse history (containing reference to the kitchen) combined with ontological world knowledge (the fridge is located in the spatial region occupied by the kitchen). This also involves a selection of *granularity*, since the speaker here moves away from particular objects and instead uses a 'niche'-like functional categorization (Bateman and Farrar, 2004). These categories are like physical objects in that they have a location but are also functionally-motivated for specific purposes and entities.

It is not, however, up to the linguistic semantics to provide the necessary constraints that fix the contextualized interpretation of these expressions. The mapping that we are seeking between the linguistic and the conceptual ontological descriptions consists of, on the one hand, particular spatial relations and entities in the linguistic semantics and, on the other, particular world-anchored spatial quality regions in the conceptual ontology. For this, we must turn to the conceptual ontology and its characterization of the room to find the other 'half' of the mapping pair.

We obtain this by examining the abstract geometric structures associated with each GUM spatial category, adopting this as a 'plug-and-play' spatial ontology module and anchoring the physical entities picked out as relatum, origin, and locatum into the positions or regions specified. In addition, in order to be anchored into the conceptual ontological description, the linguistically-free variables of orientation and origin need to be fixed: it is primarily the record of these bindings that constitutes the dynamic state of the mapping established through the discourse. We illustrate this with respect to the first five spatial expressions of the dialogue extract; how this proceeds for the rest of the extract should then be relatively clear.

First, we have the semantic spatial configuration:

NonprojectionAxis: opposite (locatum: desk, relatum: self)

The term 'gegenüber' ('opposite') presupposes a view direction or an entity with an intrinsic front. In our case, this is directly given by the phrase 'von hier aus', implying that the speaker's front and point of view serves as the required basis for direction. The term furthermore conveys a sense of at least a certain amount of distance. This then anchors both a 'reference entity' and an orientation axis through that entity as required for a 'front'/'back' abstract geometry (cf. Figure 5). Somewhere in that front half-plane, we find the location of a further physical endurant, labelled as a desk. 'Opposite' also seems more approriate when the locatum is similarly 'facing back' to the relatum and so we can construe the spatial relationship equally in terms of two symmetric TopographicDirectionals, each pointing towards the other. This constructs then two physical endurants in the conceptual ontology whose location qualities are related in the way described: i.e., in terms of twinned opposite directions requiring commitment to an orientation of the space according to the 'front' direction of the wheelchair/speaker.

Second, we have the semantic spatial configuration:

LeftProjectionExt (locatum: computer, relatum: desk)

The abstract geometry activitated by this GUM category was also shown in Figure 5 and again our task in contextualizing the expression is to adopt this abstract geometry as our characterization of space and to anchor physical endurants appropriately. This time the half-plane axis passes through the desk which we located in the previous step. We do not yet know, however, what the orientation of this axis is because, as usual, the origin is not specified in the linguistic form. There is, then, a decision to make. There are two most likely entities that could impose orientation: the speaker and the desk itself, since both are 'oriented entities'. Exploring a range of discoursal and lexicogrammatical clues for constraining this decision would lead directly to concrete empirical questions. In order to 'slot in' this additional slice of spatial description to the developing model of the situation, however, a commitment to origin and orientation must again be made. For the present case, a continuation of the origin of the previous expression, i.e., 'self', would be quite likely. This gives an orientation 'away' from the speaker's location and identifies a half-plane for the position of the physical endurant of the computer. At this point, we might also add in we could also add in expectations from the task type of scene description that the actual position should lie near the orthogonal axis: i.e., towards the canonical left/right direction, further constraining the direction in which the relatum lies. Note that now we have extended our mapping by a further step in which three entities are related and one of these has provided an oriented origin for each of the descriptions.

Next, we have the semantic spatial configuration:

HorizontalProjectionExt (locatum: kitchen, relatum: computer)

Like some other terms occurring in our discourse extract, 'neben' ('next' or 'beside') does not involve an underlying reference system, but nevertheless shares some ontological aspects with projective terms. 'Neben' presupposes a spatial—namely, the lateral—axis; however, it is underspecified with respect to side (left or right). Therefore, it is sometimes treated as a superordinate term subsuming 'left' and 'right'. As a consequence of the underdeterminacy many complications involved in interpreting underlying reference frames do not come into play in the interpretation of 'neben': projective terms can be interpreted in directly contradictory ways, depending on perspective; this is not possible with 'neben' since this term already comprises the two spatial opposites. Additionally, it conveys a notion of nextness or close spatial vicinity. The remaining problem then is to resolve the underdeterminacy using cues from the discourse context. Often, also in the present extract, 'neben' (and its deictic counterpart 'daneben', i.e., 'next to that') is directly syntactically combined with further spatial information resolving the underdeterminacy. This spatial perspective then adds qualitative distance to the developing 'theory' of the room's layout.

Then, we have the semantic spatial configuration:

ExternalConnection (locatum: kitchen, relatum: wall)

The preposition 'an' ('at') serves to convey spatial vicinity without specifying exactly how close the locatum is supposed to be to the relatum. In the present case, 'an der Wand' ('at the wall') may convey a spatial relation of contact, but in other cases, 'at' conveys coincidence or disjunctness. Therefore, the later utterance (6) 'da is' mein Herd' (there is my stove) could be interpreted as also conveying the spatial relation 'at' without explicit mention of 'an', since the same spatial concept is conveyed by the term 'there' which deictically indicates a previously mentioned location as relatum. Adding this into our growing conceptual specification of the spatial region being described then brings region connectedness into the account.

And finally, we have the semantic spatial configuration:

RightProjectionInt (locatum: fridge, relatum: kitchen)

The abstract geometry of this category is also shown in Figure 5. When anchoring the entities identified in the discourse to this structure, the region that is intersected with the 'right/left' half-plane is provided by the spatial location of the physical endurant 'kitchen', while the orientation of that region again needs to be determined in order to anchor the specification into the conceptual ontology spatial description. In this case, however, kitchen's are not generally oriented entities and so there is no potential contribution. We still have the speaker's orientation to choose and it is partly this very consistent adoption of origin throughout the discourse extract that makes it so straightfoward despite the multiple reference frames that are taken as the tour around the room progresses.

It is not necessary then that a given discourse (or even, as we see, a given utterance) constrains itself to a single spatial description strategy. In the current discourse extract, there are frequent changes of relata, switching from object-to-object descriptions to references to the room's internal features (corner, wall) to vague pointers such as 'there', and to using the speaker's own position as a relatum. We also find instances of fundamentally different spatial concepts such as 'between 12 and 3 o'clock'. Frequent self-corrections and reformulations reinforce this constant change of description concept. This phenomenon is referred to as **mixing perspectives** by Tversky et al. (1999). Such results lead to the conclusion that speakers may not draw on a clearly specified concept of any kind of reference frame, but rather, they use fragments of spatial concepts and roles to identify spatial relations. These fragments then need to be seen in relation to the current spatial scene by the interlocutor, whose task is to identify the most probable interpretation in the light of all available information. In the terms developed here, this interpretation is provided by an explicit binding of abstract geometric space, origins and perspectives that allow the linguistically induced assertions to be inserted as partial 'theories' into the conceptual ontological description. The record of these bindings constitutes dynamic mappings maintained and developed through the discourse.

4.2 Discourse-conditioned anchoring of inter-ontology mappings

Whereas we have sketched how the linguistic spatial categories of GUM isolate abstract geometric structures for filling out a growing conceptual ontological representation, we need to go further to establish that the two levels of representation, and the division of labour that they suggest, is necessary for dealing with spatial discourse. To show this, we consider a further discourse extract, R005, which adopts a completely different strategy for precisely the same task and the same room layout. This makes it very clear that the flexibility supported by mapping between ontologies is required and cannot be shortcut with pre-specified interpretations.

Dialogue R005 German

1	von der Ausgangsposition gerade, - aus, mit nur einem ganz, - kleinen Tick nach rechts, – fährt man direkt auf den Schreibtisch zu ,	*from the starting position going straight and just a tick to the right you go directly towards the desk*
2	zwei Meter links vom Schreibtisch ist der Computertisch? (1)	*two meters left of the desk is the computer table*
3	ähm, - vom Computertisch ein Stück zurück setzen, –	*er, back up a bit from the computer table*
4	nach rechts fahrn, - ziemlich weit kommt man in den Wohnbereich wo links der Fernseher, – und rechts, - das Sofa mit dem Sessel und dem Couchtisch is', (1)	*driving to the right, quite far, you come into the living room area where the TV set is on the left, and the sofa with the armchair and the coffee table on the right*
5	wenn man sich dann wieder um hunderachtzig Grad dreht, - is' links von einem der Esstisch mit der Kopfseite, –	*when you then turn around again 180 degrees, then left of you is the front side of the dining table*
6	ähm, – davon, - rechts is' der andere freie Platz am Tisch? (2)	*er, right of that is the other empty place at the table*
7	ähm vom Tisch aus geradeaus mit ein bisschen nach rechts kommt man in die Küche, –	*er, straight one from the table, a little to the right, you come into the kitchen*
8	wo links der Herd, und rechts der Kühlschrank is'.	*where the stove is left, and the fridge is right.*

Here we see the speaker employing not only spatial localizations as seen in the previous discourse but also explicit spatial *movements*—although it is important to stress that these movements were not actual movements: the speaker was sitting in the same position in the wheelchair as indicated in Figure 6 throughout the entire fragment shown here. The movements are therefore imaginary (and probably as a consequence, subject to slight errors). The linguistic semantics given by the Generalized Upper Model for this sequence of descriptions is then as shown in Table 2; here we have in addition to the spatial locatings found in the previous example, a further set of specifications drawing on the GUM area of spatial movement;

Locatings:

utterance	locatum	relatum	GUM-category
2	computer table	desk	LeftProjectionExt
4	TV	living area	LeftProjectionInt
4	{sofa, armchair, coffee table}	living area	RightProjectionInt
5	dining table	self	LeftProjectionExt
6	free place	dining table	RightProjectionExt
8	stove	kitchen area	LeftProjectionInt
8	fridge	kitchen area	RightProjectionInt

Movements:

utterance	source	target	GUM-direction
1	starting position	desk	TopographicDirectional
3	computer table	–	BackProjectionExt
4	–	living area	RightProjectionExt
5	–	–	TopographicDirectional: 180^o
7	dining table	–	TopographicDirectional

Table 2: Linguistic ontology characterization of the spatial content of dialogue R005

space precludes a detailed description and so we focus primarily on the *direction* specified also in terms of the GUM spatial categories given above. The semantics of these categories in the context of movement is quite uniform—i.e., motion is considered so that the moved entity changes from its former position to that of the region identified by the GUM spatial category. This enables us to use precisely the same abstract geometric categories for both the movement and non-movement cases.

We see in this specification that the linguistic semantics draws on linguistic ontological classes that involve some completely different space region specializations to those of the former extract. In particular, there are notions of routes and directions and within those measurable turning angles. Also, there are several instances of internal reference frames (utterances 4 and 8) using abstract areas, i.e., the niches of the living and kitchen areas, as a relatum that (functionally) contain a number of locata. Nevertheless, the link between the linguistic specification and a contextualized interpretation proceeds in exactly the same way as illustrated for the previous dialogue extract. What is different is that some further, more route-like, abstract geometries are added in to the multiple perspectives anchored in the conceptual ontology view.

The requirement that we relate the linguistic semantic specification to the constructed conceptual specification *flexibly* is shown particularly in this example by utterance (5). This gives us a semantics of a spatial LeftProjectionExt relationship with the dining table as locatum and 'self' (i.e., the imagined position of the speaker at that point in the description) as the relatum. Contrasting this with the semantic specification given for utterance (10) in extract R020 above, we see that there it was stated that there is a spatial RightProjectionExt between the dining table locatum and 'self'. In other words, we have the stereotypical situation of 'your left or mine?' but across two distinct interactions. And, within each interaction, there was little doubt as to the correct interpretation: in both cases the discourse context sets up a likely range of mappings that positions the dining table in the same physical position despite the opposite linguistic spatial semantic specifications given. This is managed by the bindings that are established during anchoring of linguistic spatial perspectives into the conceptual ontology description being quite different at that point in the discourse. Interlocutors here would, accordingly, not be any doubt about which anchoring is probably intended.

Whereas we have assumed in the discussion so far that the linguistic ontology abstract geometrical structures fit readily with the conceptual ontology description being constructed, in the following final example from this corpus we show that the linguistic semantics and an appropriate spatial anchoring are not always so easy to align. In discourse R038, the strategy employed is again different.

Dialogue R038 German

1	ähm, (1) ähm, (1) Tisch, -	*er, er Table*
2	ähm, - rechts ein Sessel, (1)	*er, right an armchair*
3	auf dem Tisch steht, die Kerze und, da liegt die Fernbedienung links davon is' der Fernseher, (2)	*on the table there's the candle, and the remote control and left from there is the TV*
4	rechts vom Sessel is', – ein anderer Tisch mit, äh, - dem Teller der Tasse und dem Löffel, und der Pflanze? (1)	*right of the armchair there's another table with, er, the plate, the cup, and the spoon, and the plant*
5	ähm, (4)	*er*
6	geradeaus und ein Stück nach rechts is' der Tisch mit der Lampe und dem Tacker und dem Lineal und dem Stift? (1)	*straight on and a bit to the right is the table with the lamp and the stapler and the ruler and the pen*
7	ähm, (1) ein Stück weiter rechts, - äh, (1) ähm, - ein, ähm , - anderer Computer, (1)	*er, a bit further right, er, er, a, er, another computer*
8	links, – is' ähm, der Tisch mit dem Monitor der Tastatur und der Maus und dem Computer, (1)	*left, is, er, the table with the monitor, the keyboard, and the mouse and the computer*
9	ähm, (4) ähm, m, links davon is' die Herdplatte mit dem Topf, (1)	*er, er, m, left of that is the stove with the pot*
10	ähm, – ähm, - ganz links von mir is' die ähm, (1) Thermoskanne mit den,, Keksen, und der Tisch (2)	*er, er, leftmost of me is the, er, thermos flask with the biscuits, and the table*

In this extract the linguistic semantics contains relatively little spatial information and the relationships expressed correspond only loosely to the physical layout. There appear to be three main spatial 'chunks' established, all internally organized along a left-right axis. The first includes the TV, the table and armchair and the dining table; the second includes the writing desk, computer table and the stove; and the third includes the speaker and the kitchen table. Here it is also interesting that in a previous phase of the experiment, where the subject was driving the wheelchair around the layout, the verbal descriptions produced by this particular speaker only used single object designations with high granularity (e.g., stapler, mouse, cakes) and virtually no spatial information at all.

The information presented is nevertheless correct, although it cannot be mapped onto a simple rectangular space. Doing so would produce a 'projection' of the room as shown on the left of Figure 7. In fact, the subject appears more to be scanning around the room and so the appropriate space region quality space would relate left-right more to relative direction around a *circular* viewing space than to straight half-planes. This kind of use shows dramatically that the relationship between the conceptual ontology and the linguistic ontology needs to operate at a sufficiently high level of abstraction so that such carry-over of use is natural. Restricting left-right to an ordering relation that may be mapped to any spatial configuration as necessary is one way of doing this and is the the naturally preferred approach according

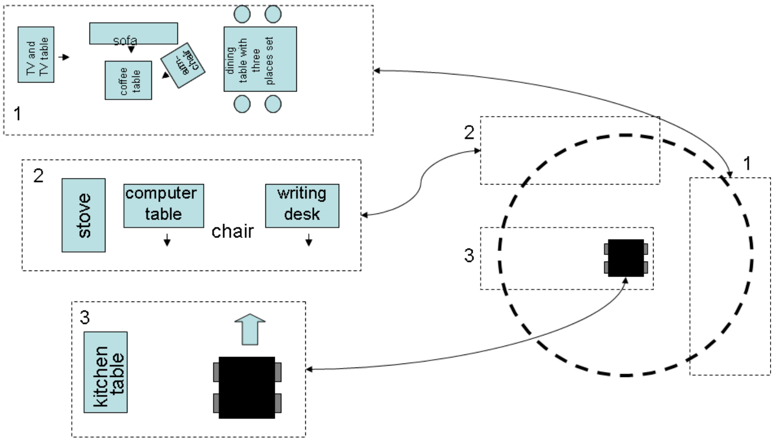

Figure 7: Spatial description and projection onto circular scan around room

to our separation of ontological levels.

Within a single discourse, then, there may be different anchorings of spatial perspectives in force than might be required for another discourse *even if the same linguistic terms and linguistic semantics are being employed*. The argument follows that, while the level of detail required for complex spatial reasoning is fairly high and calls for detailed formalizations of space at the conceptual ontological level, the corresponding level of linguistic semantic detail needed for interpretation is relatively low. Distinguishing the work performed by the different ontological layers provides a simplification of the task overall and encourages flexible, but constrained, interpretation as required.

5 Conclusions and future outlook

In this paper, we have seen how we can employ a two-level approach to ontological modelling in order to separate out some of the component tasks in semantic description for natural utterances. The two-level approach gives the necessary flexibility to describe what is semantically common to a variety of uses of spatial expressions while still allowing context-specific interpretations. The application of ontological modelling allows us to impose strict modelling conditions on both levels involved in the process. Building on the current state of the art in formal ontology, we have been able to draw on a progression in the sophistication of the relationship assumed necessary between the distinct kinds of information deployed during interaction. In terms of formal ontology and relationships required between ontological modules, this progression can be seen as one of moving from subsumption (as in the Penman Upper Model and most current approaches to the use of ontology for discourse processing) to import/export, underspecification and simple mappings/alignments (as pursued in the context of the Semantic Web) and finally to theory morphisms, whereby distinct abstract theories of space are related to one another and entities and relations are anchored across those models.

In our account, the linguistic ontology determines precisely which entities are to be anchored in terms of specific spatial roles and which abstract spatial models are to be applied as specifications of abstract space regions. The conceptual ontology determines types of object and their associated physical location possibilities, physical behaviour, etc. and the formal binding between those objects and abstract space regions. The mapping *between* these then sets discourse entities and objects into relation, thereby providing information about their physical locations. Distinct perspectives on space are combined via theory morphisms across the perspectives anchored by the related discourse entity and object pairs as constants of the related theories. This provides a generic mechanism for capturing spatial interpretation in discourse that allows selected linguistic form to signal the range of potential space models that should be applied for interpretation without limiting the flexibility of their application.

Our results are relevant for a number of ongoing discussions as well as for computational instantiations of the discourse interpretation process. The formal specification of ontologies and mappings allows a range of implementations, which we are actively following up in our human-robot-interaction work (cf. Ross et al., 2005). The ontologies are specified in standard knowledge representation languages and ontology specification formalisms, while the mappings between them are expressed as explicit links between the related modules. Moreover, since the linguistic ontology is specifically constructed to maintain a clear link with linguistic form, we are working with natural language generation and analysis components who input and output forms are semantic specifications directly employing the GUM categories. This therefore provides a ready link to the next stage of the process, the one which we have focused on this paper, whereby these specifications are anchored into the conceptual ontology interpretation. In this way our account contributes significantly to computational discourse processing systems.

Our results are also relevant to ongoing discussions of alignment in spatial dialogue, as established by Schober (1993, 1998). Early results on lexical and syntactic alignment have been followed by arguments that there is also alignment of representations at the conceptual level—and, particularly relevant here, spatial reference frames. Our decomposition of the discourse process in terms of distinct ontological representations that are brought into explicit relationships during discourse allows us to focus more on precisely *what* is being aligned from the ontological perspective. Schober (1993), for example, has shown that interlocutors coordinate reference objects and origins of reference frames. We can now describe this in terms of the construction of correspondence between constants of the theories defined by the spatial quality spaces. Moreover, Carlson-Radvansky and Jiang (1998) have shown that when a reference frame has been selected on the basis of a direction such as *left*, then the effects hold for subsequent use of both *left* and *right*, i.e., the entire dimension is involved. Now, as we have shown in the previous section, we can consider this more closely—what does it mean to say that an 'entire dimension' is involved? Only if this entity (e.g., an axis, half-plane, etc.) is available in the space quality region being used can it be available for selection.

Finally, the precise conditions under which such mappings are established, as well as their precise contents (i.e., what constitutes form, which aspects of form, which aspects of meaning, etc.?), are still unclear in many respects and there is a considerable need for more focused and detailed research from both linguistic and psychological angles. Our own contribution to this endeavor within the context of our proposed ontology-based architecture provides a formally well-specified starting point for framing the detailed empirical research questions necessary.

Acknowledgements

The work reported in this paper was carried out with the financial support of the Deutsche Forschungsgemeinschaft (DFG) within the Transregional Collaborative Research Centre SFB/TR 8. We thank the members of the OntoSpace, SharC and SPIN projects of the center and particularly Joana Hois and Robert Ross for their contributions to the development of the linguistic ontology.

References

Bateman, J., Borgo, S., Lüttich, K., Masolo, C., and Mossakowski, T. (submitted). Ontological modularity and spatial diversity.

Bateman, J., Farrar, S., Hois, J., Tenbrink, T., and Vierhuff, T. (2006). Augmenting the Generalized Upper Model for spatial language. SFB/TR8 internal report, Collaborative Research Center for Spatial Cognition, University of Bremen, Germany.

Bateman, J. A. (to appear). Linguistic interaction as ontological mediation. In Zaefferer, D. and Schalley, A., editors, *Ontolinguistics. How ontological status shapes the linguistic coding of concepts*. Mouton de Gruyter.

Bateman, J. A. and Farrar, S. (2004). Towards a generic foundation for spatial ontology. In Varzi, A. C. and Vieu, L., editors, *Formal Ontology in Information Systems: Proceedings of the Third International Conference on Formal Ontology in Information Systems (FOIS-2004)*, pages 237–248, Amsterdam. IOS Press.

Bateman, J. A., Henschel, R., and Rinaldi, F. (1995). Generalized upper model 2.0: documentation. Technical report, GMD/Institut für Integrierte Publikations- und Informationssysteme, Darmstadt, Germany.

Beun, R.-J., van Eijk, R., and Prüst, H. (2004). Ontological feedback in multiagent systems. In Jennings, N. R., Sierra, C., Sonenberg, L., and Tambe, M., editors, *Proceedings of the 3rd international Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004)*, pages 110–117. ACM Press.

Blackburn, P., Gardent, C., and de Rijke, M. (1996). Rich ontologies for tense and aspect. In Seligman, J. and Westerstahl, D., editors, *Logic, Language and Computation*. CSLI Publications, Stanford.

Carlson-Radvansky, L. and Jiang, Y. (1998). Inhibition accompanies reference frame selection. *Psychological Science*, 9:386–391.

Carlson-Radvansky, L. A. and Logan, G. D. (1997). The influence of reference frame selection on spatial template construction. *Journal of Memory and Language*, 37:411–437.

Denis, M. (1997). The description of routes: A cognitive approach to the production of spatial discourse. *Cahiers de Psychologie Cognitive*, 16:409–458.

Eschenbach, C. (1999). Geometric structures of frames of reference and natural language semantics. *Spatial Cognition and Computation*, 1(4):329–348.

Farrar, S. and Bateman, J. (2004). General ontology baseline. SFB/TR8 internal report I1-[OntoSpace]: D1, Collaborative Research Center for Spatial Cognition, University of Bremen, Germany.

Fischer, K. (2003). Linguistic methods for investigating concepts in use. In Stolz, T. and Kolbe, K., editors, *Methodologie in der Linguistik*, pages 39–62. Peter Lang, Frankfurt a.M.

Fischer, K. (2006). *What Computer Talk Is and Is not: Human-Computer Conversation as Intercultural Communication*, volume 17 of *Linguistics – Computational Linguistics*. AQ-Verlag, Saarbrücken.

Gapp, K.-P. (1995). An empirically validated model for computing spatial relations. In *KI - Künstliche Intelligenz*, pages 245–256.

Garrod, S. C. and Sanford, A. J. (1988). Discourse models as interfaces between language and the spatial world. *Journal of Semantics*, 6:147–160.

Hayes, P. J. (1979). The naive physics manifesto. In Michie, D., editor, *Expert systems in the microelectronic age*. Edinburgh University Press, Edinburgh, Scotland.

Herrmann, T. and Deutsch, W. (1976). *Psychologie der Objektbenennung*. Hans Huber Verlag, Bern u.a.

Herskovits, A. (1986). *Language and Spatial Cognition: an interdisciplinary study of the prepositions in English*. Studies in Natural Language Processing. Cambridge University Press, London.

Hobbs, J. R. (1995). Sketch of an ontology underlying the way we talk about the world. *International Journal of Human-Computer Studies*, 43(5/6):819–830.

Klippel, A., Tenbrink, T., and Montello, D. (forthcoming). The role of structure and function in the conceptualization of directions. In van der Zee, E. and Vulchanova, M., editors, *Motion Encoding in Language and Space*. Oxford University Press, Oxford. under review.

Lankenau, A., Meyer, O., and Krieg-Brückner, B. (1998). Safety in robotics: The Bremen autonomous wheelchair. In *Proceedings of the 5th International Workshop on Advanced Motion Control (AMC98)*, pages 524–529.

Levinson, S. C. (2003). *Space in language and cognition: explorations in cognitive diversity*. Cambridge University Press, Cambridge.

Mann, W. C., Arens, Y., Matthiessen, C. M. I. M., Naberschnig, S., and Sondheimer, N. K. (1985). Janus abstraction structure — draft 2. Technical report, USC/Information Sciences Institute, Marina del Rey, California. (Circulated in draft form only.).

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., and Oltramari, A. (2003). Ontologies library (final). WonderWeb Deliverable D18, ISTC-CNR, Padova, Italy.

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A., and Schneider, L. (2002). The WonderWeb library of foundational ontologies: preliminary report. WonderWeb Deliverable D17, ISTC-CNR, Padova, Italy.

Moratz, R. and Tenbrink, T. (2006). Spatial reference in linguistic human-robot interaction: Iterative, empirically supported development of a model of projective relations. *Spatial Cognition and Computation*, 6(1):63–107.

Pickering, M. J. and Garrod, S. (2004). Towards a mechanistic psychology of dialogue. *Behavioural and Brain Sciences*, 27(2):169–190.

Ross, R., Shi, H., Vierhuff, T., Krieg-Brückner, B., and Bateman, J. (2005). Towards dialogue based shared control of navigating robots. In Freksa, C., Knauff, M., Krieg-Brückner, B., Nebel, B., and Barkowsky, T., editors, *Spatial Cognition IV: Reasoning, Action, Interaction. International Conference Spatial Cognition 2004, Frauenchiemsee, Germany, October 2004, Proceedings*, pages 478–499, Berlin, Heidelberg. Springer.

Schober, M. F. (1993). Spatial perspective taking in conversation. *Cognition*, 47:1–24.

Schober, M. J. (1998). How addressees affect spatial perspective choice in dialogue. In Olivier, P. and Gapp, K.-P., editors, *Representation and processing of spatial expressions*, pages 231–246. Lawrence Erlbaum Associates, Mahwah, New Jersey.

Tenbrink, T. (2005a). Identifying objects in English and German: Empirical investigations of spatial contrastive reference. In *WoSLaD Workshop on Spatial Language and Dialogue, October 23-25*.

Tenbrink, T. (2005b). Identifying objects on the basis of spatial contrast: an empirical study. In Freksa, C., Knauff, M., Krieg-Brückner, B., Nebel, B., and Barkowsky, T., editors, *Spatial Cognition IV: Reasoning, Action, Interaction. International Conference Spatial Cognition 2004, Frauenchiemsee, Germany, October 2004, Proceedings*, pages 124–146, Berlin, Heidelberg. Springer.

Tenbrink, T. (2005c). *Localising objects and events: Discoursal applicability conditions for spatiotemporal expressions in English and German. Dissertation.* PhD thesis, University of Bremen, FB10 Linguistics and Literature, Bremen.

Tenbrink, T. and Moratz, R. (2003). Group-based spatial reference in linguistic human-robot interaction. In *Proceedings of EuroCogSci 2003: The European Cognitive Science Conference*, pages 325–330.

Tenbrink, T., Shi, H., and Fischer, K. (2006). Route instruction dialogues with a robotic wheelchair. In *Proc. BranDial 2006: The 10th Workshop on the Semantics and Pragmatics of Dialogue. University of Potsdam, Germany; September 11th-13th 2006.*

Tenbrink, T., Winterboer, A., and Moratz, R. (forthcoming). Motion instructions by dynamic projective terms: Linguistic considerations and experiences in human-robot interaction. In van der Zee, E. and Vulchanova, M., editors, *Motion Encoding in Language and Space.* Oxford University Press, Oxford. under review.

Tversky, B., Lee, P., and Mainwaring, S. (1999). Why do speakers mix perspectives? *Spatial cognition and computation*, 1:399–412.

Vorwerg, C. and Tenbrink, T. (2006). Discourse factors influencing spatial descriptions in English and German. In *International Conference Spatial Cognition 2006, Bremen.*

Wölfl, S. and Mossakowski, T. (2005). CASL specifications of qualitative calculi. In Cohn, A. G. and Mark, D. M., editors, *Proceedings of Spatial Information Theory: International Conference, COSIT 2005*, number 3693 in LNCS, pages 200–217, Berlin. Springer.

Zimmer, H. D., Speiser, H. R., Baus, J., Blocher, A., and Stopp, E. (1998). The use of locative expressions in dependence of the spatial relation between target and reference object in two-dimensional layouts. In *Spatial Cognition I - An interdisciplinary approach to representing and processing spatial knowledge*, pages 223–240. Springer.