On the Role of Conceptual and Linguistic Ontologies in Spoken Dialogue Systems

Scott Farrar, Thora Tenbrink, John Bateman, Robert J. Ross

May 10, 2005

Abstract

We report on the role of well-formed conceptual and linguistic ontologies in empirically grounded 'spoken dialogue systems' (SDS). In particular we use empirical results from spatial dialogues in German to argue for the strict separation of linguistically motivated knowledge from non-linguistic, domain concerns. We motivate our arguments with a number of examples relevant to the language generation task, and show how a well-defined separation of linguistic and domain concerns can be effected in a practical SDS.

1 Introduction

The development of robust 'spoken dialogue systems' (SDS) requires the convergence of empirical, formal, and practical design principles. We argue that empirical results from human-robot interaction tasks can be applied to create a formal model of dialogic interaction that can then be implemented on real robotics systems. Furthermore, the formal modeling stands to benefit form recent progress in applied formal ontology, both from a linguistic and non-linguistic point of view. Given the difficulty in coordinating the empirical, formal, and practical consideration, we focus on the specific domain of space. The main task then is to create an SDS for spatial language, in particular, one that can be implemented as part of a mobile robot.

The paper is organized as follows. In Section 2 we present background concerning the nature of spoken dialogues, our approach to NLG, and ontologies. Based on our empirical results, we present in Section 3 the argument for a two-level approach and situate our results by discussing an implementation of a practical SDS for the spatial domain. Finally in Section 4 we turn to the formalization of the relationship between non-linguistic, conceptual knowledge and linguistic knowledge.

2 Background

In the following section we lay the background for our treatment of language generation about space. First we focus on the nature of dialogues about space and identify some of the key tasks and strategies used. Second, we give a brief overview of our approach to NLG in general and highlight some basic assumptions. Finally, we describe two kinds of ontologies, one motivated by linguistic means and one motivated according to particular non-linguistic conceptualizations of space.

2.1 Spatial Dialogues

Our investigation of spatial dialogue emerges from empirical research aimed at establishing how human users speak to robots about spatial situations. In any dialogue concerning space, a number of generation-specific tasks can be identified. (Since our focus is on language generation, we will have little to say regarding language understanding except for how the understanding component is implemented in our dialogue system, but see Section 3.1). The generation tasks are: (i) to describe configurations of objects, or 'scene description', (ii) to refer to an object by identifying its spatial relationship to one or more other objects, or 'referential identification', and (iii) to give 'route descriptions' (Tenbrink & Klippel 2005). In our experiments, route descriptions are the instructions given to robots to follow a certain route and thus arrive at a goal location. Describing a spatial scene is the task of establishing where various objects are located with respect to one another. Identifying an object via spatial reference occurs when one object needs to be singled out in a scenario where several competing objects are present, such as when the robot is required to move toward one of several similar objects.

What makes spatial dialogue particularly interesting, and challenging to implement in robotic systems, is that human users employ a number of very different strategies in order to complete the above tasks. In spatial scene descriptions, for instance, speakers typically use locative **dimensional** terms like *links/rechts* in (1) and *vor/hinter* in (2). Locative dimensional terms are those that locate objects with respect to one another on a specific spatial dimension or axis (i.e., the vertical, front-back, or lateral dimesions) (Tenbrink 2005*b*, p.1).

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

(2) Das Auto ist vor dem Haus.
'The car is in front of the house.'

Scene description. The application of locative dimensional terms in spatial scene descriptions like these depends on the particular spatial relationship between the object to be identified, called the **referent**, and another object against which the spatial relation is formulated, called the **relatum**. Unmodified dimensional terms may be used when the referent is positioned directly on a spatial axis with respect to an underlying **reference system**. (We will have more to say about reference systems at the end of this section.) With increasing angles within the spatial scene, modifications of the spatial term and combinations of several terms become increasingly likely. These effects have been described in terms of **spatial templates**, for example, by Carlson-Radvansky & Logan (1997). Apart from locative dimensional terms, spatial scenes can be described using topological (e.g., *in*/'in') or path-related terms (e.g., *durch*/'across'), expressions denoting distance or describing in-between relations, and the like (Tenbrink 2005*b*).

Referential Identification. In tasks requiring referential identification, speakers may refer to the intended object either by using these same kinds of spatial expressions, as described for the scene description task, e.g., *zum linken Karton*/'to the left box', or they may use discriminating features such as color, shape, and size. In some cases, even the class type of the object may be sufficient if it is discriminative. Thus, referential identification involves the use of similar terms but in ways different from those used in other kinds of tasks, since an object needs to be singled out on the basis of contrast to other objects present (Herrmann & Deutsch 1976). Such a contrast may be spatial or involve object features. If locative dimensional terms are used, they only need to be modified if other objects can also be described by the same spatial term. Otherwise, spatial templates do not necessarily come into play (Tenbrink 2005*a*).

Route description. If the main task is to direct an interlocutor to move towards a goal, i.e., a route description, speakers may refer directly to the goal itself as in (3), using referential identification as just described, and leaving path information unspecified. By contrast speakers may refer to the path itself and not the goal entity, as in as in (4). Full descriptions contain information about path and goal, as in (5).

(3) Gehen Sie zu dem Quadrat ganz hinten rechts.
'Go to the square all the way back and to the right.'

(4) Gehe geradeaus, leicht nach rechts.
'Go straight ahead, slightly to the right.'

(5) Gehen Sie geradeaus an den Quadraten rechts von Ihnen vorbei. Es ist dann das hinterste Quadrat.
'Go straight ahead past the squares to the right of you. It is then the backmost square.'

Full descriptions may contain many kinds of additional information, such as detailed directional information and reference to further entities in sub-goals such as obstacles or landmarks for orientation. This kind of description is typically dealt with in terms of route description tasks; often, this kind of task involves a structured environment such as a street network. A broad range of additional factors come into play in this area, such as information about decision points (Denis 1997), spatial chunking processes (Klippel, Tappe & Habel 2003), identification of landmarks and other phenomena involved in cognitive wayfinding processes (Siegel & White 1975), and the interaction with previous knowledge or that derived from other kinds of representations, such as maps (Klippel 2003). The two other kinds of spatial task described above may both come into play in achieving the main route-description task, i.e., describing the spatial scene as such or identifying an object along the way. How much, and which kind of, information is actually conveyed depends on the choice of **level of granularity** (Bateman & Farrar 2004), or simply how much detail is needed for the task.

Finally, all usages of locative dimensional terms involve the conceptual factors of **perspective** and, as mentioned earlier, **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 in (6), this strategy was rarely encountered in our empirical research. Instead speakers typically only refer explicitly to the relatum as in 1 above, although this kind of information may also be omitted.

(6) Die Tür ist von dir aus gesehen links.
'The door is to the left from your point of view.'

Reference systems, on the other hand, involve the assignment of spatial roles to various objects in the spatial scene, where the roles include **referent**, **relatum**, and **origin**. The referent is the object of primary interest in the spatial generation task, that which is to be identified. The relatum is the object against which the spatial relation is formulated. And the origin is the position defining the perspective, as just mentioned. There are three types of reference systems used in human cognition, and thus reflected in natural language: intrinsic, relative, and absolute reference systems. In the case of **intrinsic** reference systems, the origin is conflated with the relatum (Levinson 2003), as becomes clear when the examples 1 and 6 are compared: both utterances are suitable for describing the same spatial concept. In **relative** reference systems, the relatum differs from the origin, as in example (7):

(7) Der Ball liegt von dir aus gesehen links vom Tisch.
'The ball lies to the left of the table from your point of view.'

In **absolute** reference systems, an established frame of reference is used, such as the compass coordinates or recognized geographical feature, to establish the spatial relation. In such a reference system, there is no way of labeling a relatum or origin as in the other kinds of reference systems. The present discussion is simplified somewhat, but there are a number of additional conceptual options for establishing reference systems, detailed in (Tenbrink 2005*b*). The main point to be emphasized, then, is that a primary conceptual task during a spatial dialogue is the negotiation of a reference system, or the dynamic assignment of the various roles. The main problem here lies in the fact that the linguistic surface form does not directly reflect underlying reference systems, except in the case of absolute reference systems, which use specific linguistic expressions such as *north* or *south*.

2.2 Approach to NLG

Our main inspiration from a linguistic point of view is Systemic Functional Grammar (SFG) (Halliday 1985, Halliday & Matthiessen 1994, Halliday & Matthiessen 1999). SFG will not be discussed in depth. However, one aspect is particularly important concerning NLG: the notion of a **metafunction**. A metafunction can be described as a particular mode, facet, or layer of meaning. SFG holds that there are three metafunctions encoded in the sentences of natural language: a **textual**, an **interpersonal**, and an **ideational** metafunction. The textual metafunction captures the meaning of the clause as 'message', or how it is used to construct a text. The textual metafunction is manifested by the theme-rheme and information structure of the grammar. The interpersonal metafunction captures the meaning of the clause as 'interaction', or how it is used to act in a discourse. The interpersonal metafunction is associated with the mood element of the grammar. Finally, the ideational metafunction captures the meaning of the clause as 'experience', or the propositional content of the sentence. The ideational metafunction can be seen, for example, in a language's transitivity system and reflects the way the grammars classify and organize the world. The current work focuses on the last metafunction, that of ideation, and how the corresponding knowledge component, the 'ideation base', can be used in implementations of SFG in the context of the Semantic Web.

One of the largest applications of Systemic Functional Grammar to NLG is the The KOMET-Penman Multilingual Development Environment (KPML) system, a large-scale grammar engineering environment with generation capabilities (Bateman 1994, Bateman 1995). The generation component (Matthiessen & Bateman 1991) extends the sentence generation component of the Penman text generation system (Mann 1983, Mann & Matthiessen 1983) and uses as input the Sentence Planning Language (Kasper 1989).

One of the major knowledge components of the PENMAN system was the PENMAN Upper Model (Mann, Arens, Matthiessen, Naberschnig & Sondheimer 1985, Bateman, Kasper, Moore & Whitney 1990). Drawing on grammars of English and German, the PENMAN Upper Model was transformed into the Merged Upper Model (Henschel 1993, Henschel & Bateman 1994) for use in the KPML system. The next step was to provide more linguistic coverage, both in terms of the generation ability in a given language, but also in various other languages, e.g., Italian (Bateman, Magnini & Rinaldi 1994) and to bring the Merged Upper Model more in line with the systemic work of Halliday and Matthiessen (Halliday & Matthiessen 1999). This effort resulted in the Generalized Upper Model (Bateman, Henschel & Rinaldi 1995). The Generalized Upper Model has also been adapted to Spanish (del Socorro Bernados Galindo & Aguado de Ceo 2001). We are currently augmenting the Generalized Upper Model with a rich collection of spatial concepts in order for application in robust spoken dialogue system.

2.3 Linguistic and Conceptual Ontologies

One thread of research for the Spoken Dialogue System is the use of conceptual and linguistic ontologies, respectively for general, non-linguistic reasoning tasks and natural language processing. By conceptual ontologies we refer to such efforts as the Suggested Upper Merged Ontology (SUMO) (Niles & Pease 2001) one of three starter documents currently under consideration by the IEEE working group for a Standard Upper Ontology (SUO), or the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) (Masolo, Borgo, Gangemi, Guarino & Oltramari 2003) a part of the WonderWeb project[1]. Conceptual ontologies are constructed such that categories are intended to be language-neutral, motivated according to cognitively or other non-linguistic criteria. Conceptual ontologies contrast with linguistically motivated ontologies, such as the Generalized Upper Model (GUM), whose categories are motivated strictly according to the grammar of natural language. We argue that the variability of human language is too great, except perhaps in very limited domains, simply to map words and phrases onto the conceptual level using a conceptually motivated ontology (Bateman, Fischer & Tenbrink 2003). Instead, we advocate keeping the conceptual and linguistic levels separate in dialogue systems, using the former for interaction with the non-linguistic components and the latter for interaction with the NLP components. Our approach embraces that of 'two-level' semantics after Bierwisch (1982) and others.

3 Rationale for the Separation of Knowledge Sorts

In this section we provide arguments for a strict separation of knowledge sorts, referred to in the linguistics literature as two-level semantics. We argue that evidence for the separation emerges from our analysis of empirical data collected during dialogues concerning spatial tasks. After some methodological preliminaries, we argue that establishing a reference system and perspective requires the separation of knowledge sorts.

3.1 Methodological Preliminaries

We have applied the arguments above to the development of a spoken dialogue system in the robotics domain (Krieg-Brückner, Shi & Ross 2004). In this implementation, depicted in Figure 1, domain components express and exchange spatial and task knowledge through purely conceptual representations (grounding in the conceptual ontology). For the output channel, strategic language generation decisions result in the generation and dispatching of conceptual representations to dialogue management. At this point, relevant content for linguistic output is identified, before ontological mappings are applied to build semantic representations for the tactical generator KPML (Bateman 1997). Thus, the non-linguistic, conceptual representations are transformed into a representation whose categories derive from a linguistically motivated source – the linguistic ontology. Similarly, in the reverse, a linguistic-to-conceptual mapping is performed for the input channel – this has been presented elsewhere (Krieg-Brückner et al. 2004).

The architecture in Figure 1 reflects a rather strict two-level semantics. The linguistic ontology captures the 'ideational metafunction' component of the meaning of utterances, as mentioned in Section 2.2. In other words, the linguistic ontology provides the categories that are used to build a symbolic representation of the propositional content. In the current work we consider only the propositional content of the linguistic semantics. (Giving an ontological account of the other two metafunctions is future work.) The conceptual ontology on the other hand provides the language-neutral categories necessary to build a representation for space and robot control. We refer to the first level of representation as **semantic representation** and the latter as **conceptual**

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

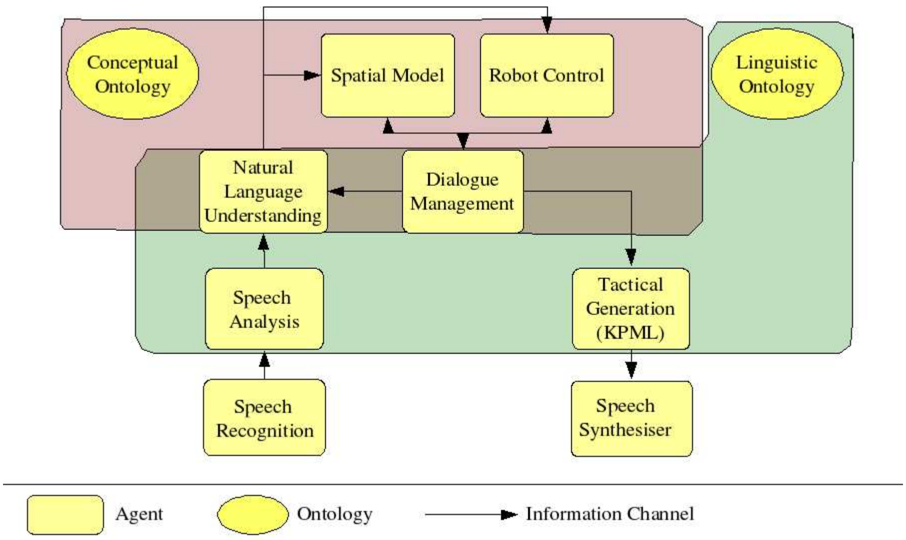

Figure 1: The SharC Spoken Dialogue System

representation. We note that categories from both ontologies are necessary for a complete symbolic representation.

3.2 The Problem of Logical Structure

The first argument in favor of a two-level approach is that having a single level of symbolic representation is incommensurate with having multiple symbolic reasoning tasks such as those required of the robot. On the one hand, there is the task of semantic representation for language. The ideal semantic representation should facilitate tasks such as language generation and understanding by providing a logical structure that maps easily onto the lexicogrammar. This means, for example, that the degree of abstraction assumed of the logical structure should match that of the lexicogrammar, since the logical structure of the semantics follows to a large degree from surface form. On the other hand, a mobile robot requires a symbolic representation also for various non-linguistic tasks, e.g., route planning and object recognition. To achieve these tasks, the logical structure of the semantic representation should assume an ontology that is compatible with the various computational calculi used for non-linguistic tasks.

Another argument is that a two-level approach places less burden on the grammar which greatly facilitates the design of generation systems. With one level of representation, complex mapping rules are required between the logical structure and the lexicogrammar. With the two-level approach, complex mapping rules are required between conceptual and semantic representation.

From a software engineering standpoint, separating the conceptual from the linguistic allows greater modularity: the mappings to the lexicogrammar (that used by the generation system) are maintained as maximally simple even though the linguistic expression of the conceptual information may become extremely varied. This is one of the points made at length in Bateman (1992) who argued for the separation of grammar building (computational linguistics) from content engineering (applied ontology). The linguistic ontology acts an intermediary between non-linguistic knowledge and the lexicogrammar. In an ontology that partitions the meaning space in a way that complies with and accommodates distinctions in the morphology and syntax, it is a much simpler task to generate grammatical sentences. As an example, consider that the preposition *in* can cover a wide range of meanings, as shown in the second line of the following:

(8) a. The chair is in the robot's way.

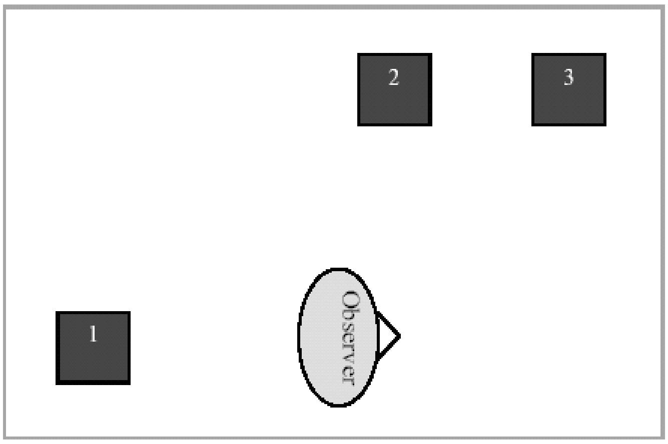

Figure 2: Conflicting Reference Systems

b. $...Chair(x) \wedge Robot(y) \wedge positionalLoc(x, y)$

(9) a. The chair is in the office.
b. $...Chair(x) \wedge Office(y) \wedge spatiallyContains(x, y)$

(10) a. There's a chair in the road.
b. $...Chair(x) \wedge Road(y) \wedge surfaceContact(x, y)$

The argument is that, while the level of detail required for complex spatial reasoning is fairly high in the above examples, the corresponding level of detail needed to produce a well formed sentence is relatively low. The scene in each example is actually conceptualized differently: as a point on a line, as a 3D object contained in a 3D object, and as a 3D object on the surface of a2D object, respectively. But all that is needed in GUM in order to generate the well-formed utterances is a relation, such as $spatialOrdering(x, y)$. Moreover, the task of representing all the meanings of such a polysemous lexical item in a single ontology would require that concepts are consistently classified along the multiple dimensions simultaneously: which complicates the formal properties of the resulting ontology considerably since exactly what may be inherited where becomes unclear.

3.3 The Problem of Reference Systems

The issue of conceptual versus linguistic knowledge is even more acute when reference systems are taken into account. We will argue that since language does not explicitly encode information about reference systems, the burden is put on the conceptual representation. Describing reference systems, then, is facilitated by keeping linguistic factors out of the conceptual description. Take, for example, the spatial situation depicted in Figure 3.3. To refer to one of the objects, the German adjective *hinteren* could be used as in (11):

(11) Geh zum hinteren Objekt.
'Go to the object in back.'

However, due to the availability of different reference systems, *hinteren* ("in back") is ambiguous and can be applied to at least two of the objects, 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 2005*b*)) reference system using the observer as origin. Thus, object 3 is situated behind object 2, at least if the view direction is shifted in parallel in the direction of the row that is suggested by the positions of object 2 and object 3. Object 1, on the other hand, is situated behind the observer and can therefore also be referred to as *hinteren*, if the underlying reference system is intrinsic, using the observer as both origin and relatum.

The problem, then, is one of negotiating reference systems. That is, the speaker must impart his or her view of the world using language, and in such a way that intentions are recoverable based on shared contextual knowledge.

4 Formalizing the Problem

In this section we present a formalization of the spatial dialogue task. In particular we focus our formalization on the task of 'scene description' in spatial dialogues, as introduced in Section 2.1. The following shows the steps in the generation process, from dialogue planning to tactical generation.

1. Choose a dialogue task (e.g., 'scene description').
2. Focus on the particular object to be described within the scene.
3. Choose the reference system to be employed in the description.
4. Assign to the objects in the scene various reference system roles.
5. Choose a perspective on the scene.
6. Select the particular spatial relation for the scene.

Consider that the particular scene to be described is that in Figure 3.3 where object 2 is a door. A relevant statement in the scene description might be as in 12.

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

That is, the goal is to refer to an object by identifying how it is ordered spatially in relation to other objects. On a purely symbolic level, we start with a conceptual representation using DOLCE. The representation requires a **location scheme**, the 4-tuple $\langle PO, RefSys, SC, FoR\rangle$ where:

- PO is the set of physical objects to be described;
- $RefSys$ is a reference system as mentioned previously;
- SC is a spatial calculus that provide relations at a certain level of detail;
- FoR is a frame of reference such that all objects in PO are 'contained within' FoR[2].

We then define a reference system as a triple $RefSys = \langle R, PO, Fn\rangle$ where:

- R is a specialized set of spatial roles: $referent, relatum, origin$;
- PO is non-empty set of physical objects;
- $RefFn$ is an assignment function such that every object in PO is assigned to a role in R.

[2]For more details of how this is done with DOLCE, refer to Bateman & Farrar (2005) for examples.

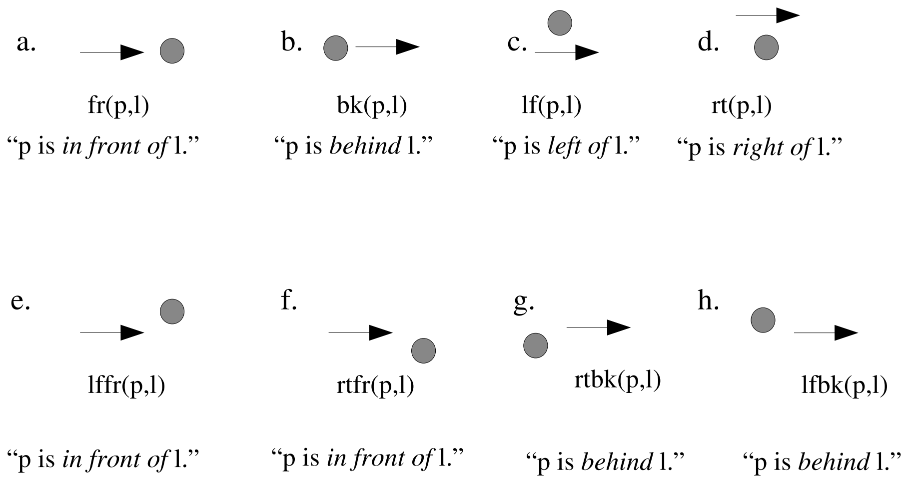

Figure 3: A simple spatial calculus for spatial ordering

The various types of reference systems can then be defined. For example, an intrinsic reference system is when the the relatum also plays the role of *origin*, and it has intrinsic properties such as a back and front.

With reference to a spatial calculus (SC), this a particular algebra that includes a number of predefined relations and supporting axioms. Depending on the task at hand, a particular, specialized spatial calculi can be employed to solve non-linguistic tasks. Consider for example a calculus useful for describing the relative positions of two objects, one with inherent orientation. Eight base relations between an oriented line (l) and a point (p) are defined for this calculus in Figure 3. Naturally the aim was to keep the number of relations admitted to the base set sufficiently small so as to support theorem proving while still covering the spatial configurations that can occur with sufficient granularity as to be useful.

The next task is to map from this non-linguistic representation to a linguistic representation using the predicates from GUM, viz. linguistic ontology. Each spatial relation expressed at the conceptual level should be mapped onto the appropriate GUM relation. From Figure 3 it can be observed that various spatial relations at the conceptual level can be generalized to a particular kind of circumstance-in-configuration from GUM. The result is an SPL such as:

```
((SP0 / SpatialPositioning
:process-in-configuration (S0 / state)
:carrier (D1 / door)
:spatial-ordering-left (H1 / hearer))
```

5 Conclusion

We advocate for the separation of linguistic from non-linguistic knowledge in Spoken Dialogue Systems. Our approach combines research from formal ontology, NLG, and empirical research in human-computer interaction tasks. We have merged these two efforts to achieve a robust dialogue system, thus eliminating the need for mixing non-linguistic, domain knowledge with that of a purely linguistic nature. Our current research focuses on developing various inter-ontological mapping algorithms such that conceptual knowledge can readily be transformed into linguistic knowledge, and vice versa. We have noted a number of advantages of the two-level approach, in particular based on our empirical research and how speakers employ, but do not express directly, reference systems in the production of spatial dialogues. Finally, having the location scheme formalized on a separate level of symbolic representation, it is then possible to incorporate various algebras for automated reasoning without altering the design of the generator.

References

Bateman, J. A. (1992), The theoretical status of ontologies in natural language processing, *in* S. Preuß & B. Schmitz, eds, 'Text Representation and Domain Modelling – ideas from linguistics and AI', KIT-Report 97, Technische Universität Berlin, Berlin, Germany, pp. 50–99. (Papers from KIT-FAST Workshop, Technical University Berlin, October 9th - 11th 1991).
URL: *http://xxx.lanl.gov/cmp-lg/9704010*

Bateman, J. A. (1994), KPML: The KOMET-Penman (Multilingual) Development Environment, Technical report, Institut für Integrierte Publikations- und Informationssysteme (IPSI), GMD, Darmstadt. Release 0.6.

Bateman, J. A. (1995), Basic technology for multilingual theory and practise: the KPML development environment, *in* 'Proceedings of the IJCAI workshop in multilingual text generation (International Joint Conference on Artificial Intelligence) 1995', Montréal, Canada, pp. 1–12.

Bateman, J. A. (1997), 'Enabling technology for multilingual natural language generation: the KPML development environment', *Journal of Natural Language Engineering* **3**(1), 15–55.

Bateman, J. A. & Farrar, S. (2005), Modelling models of robot navigation using formal spatial ontology, *in* C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel & T. Barkowsky, eds, 'Spatial Cognition IV: Reasoning, Action, Interaction. International Conference Spatial Cognition 2004, Frauenchiemsee, Germany, October 2004, Proceedings', Springer, Berlin, Heidelberg, pp. 366–389.

Bateman, J. A., Henschel, R. & Rinaldi, F. (1995), Generalized upper model 2.0: documentation, Technical report, GMD/Institut für Integrierte Publikations- und Informationssysteme, Darmstadt, Germany.
URL: *http://purl.org/net/gum2*

Bateman, J. A., Kasper, R. T., Moore, J. D. & Whitney, R. A. (1990), A general organization of knowledge for natural language processing: the PENMAN upper model, Technical report, USC/Information Sciences Institute, Marina del Rey, California.

Bateman, J. A., Magnini, B. & Rinaldi, F. (1994), The Generalized{*Italian,German,English*} *Upper Model*, *in* 'Proceedings of the ECAI94 Workshop: Comparison of Implemented Ontologies', Amsterdam.
URL: *http://www.darmstadt.gmd.de/publish/komet/papers/ecai94.ps*

Bateman, J. & Farrar, S. (2004), Spatial ontology baseline, SFB/TR8 internal report I1-[OntoSpace]: D2, Collaborative Research Center for Spatial Cognition, University of Bremen, Germany.

Bateman, J., Fischer, K. & Tenbrink, T. (2003), Why a static interpretation is not sufficient in spatial communication, *in* 'Proceedings of the European Chapter of the Association for Computational Linguistics (EACL-03) Workshop on Dialogue Systems: interaction, adaption and styles of management', Budapest, Hungary.

Bierwisch, M. (1982), Semantische und konzeptuelle Repräsentation lexikalischer Einheiten, *in* R. Ružička & W. Motsch, eds, 'Untersuchungen zur Semantik', Akademie-Verlag, Berlin, pp. 61–99.

Carlson-Radvansky, L. A. & Logan, G. D. (1997), 'The Influence of Reference Frame Selection on Spatial Template Construction', *Journal of memory and language* **37**, 411–437.

del Socorro Bernados Galindo, M. & Aguado de Ceo, G. (2001), Adapting the Generalized Upper Model to Spanish, *in* G. Angelova, K. Bontcheva, R. Mitkov, N. Nicolov & N. Nikolov, eds, 'Proceedings of the Euroconference Recent Advances in Natural Language Processing (RANLP-2001)', Tzigov, Bulgaria, pp. 103–107.

Denis, M. (1997), 'The description of routes: A cognitive approach to the production of spatial discourse', *Cahiers de Psychologie Cognitive* **16(4)**, 409–458.

Halliday, M. A. K. (1985), *An Introduction to Functional Grammar*, Edward Arnold, London. 2nd. edition, 1994.

Halliday, M. A. K. & Matthiessen, C. M. I. M. (1994), *An Introduction to Functional Grammar*, Arnold, London. 3rd. edition.

Halliday, M. A. K. & Matthiessen, C. M. I. M. (1999), *Construing experience through meaning: a language-based approach to cognition*, Cassell, London.

Henschel, R. (1993), Merging the English and the German Upper Model, Technical report, GMD/Institut für Integrierte Publikations- und Informationssysteme, Darmstadt, Germany. Appears as: *Arbeitspapiere der GMD*, **848**, June 1994. GMD, Sankt Augustin.
URL: *http://www.darmstadt.gmd.de/publish/komet/papers/merge.ps*

Henschel, R. & Bateman, J. A. (1994), The merged upper model: a linguistic ontology for German and English, *in* 'Proceedings of COLING '94', Kyoto, Japan.
URL: *http://www.darmstadt.gmd.de/publish/komet/papers/paper.ps*

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

Kasper, R. T. (1989), A flexible interface for linking applications to PENMAN's sentence generator, *in* 'Proceedings of the DARPA Workshop on Speech and Natural Language'.

Klippel, A. (2003), Wayfinding choremes, *in* W. Kuhn, M. Worboys & S. Timpf, eds, 'Spatial Information Theory: Foundations of Geographic Information Science. Conference on Spatial Information Theory (COSIT).', Lecture Notes in Computer Science, Springer; Berlin, pp. 320–334.

Klippel, A., Tappe, H. & Habel, C. (2003), Pictorial representations of routes: Chunking route segments during comprehension, *in* C. Freksa, W. Brauer, C. Habel & K. F. Wender, eds, 'Spatial Cognition III — Routes and Navigation, Human Memory and Learning, Spatial Representation and Spatial Learning', Lecture Notes in Artificial Intelligence, Springer; Berlin, pp. 11–33.

Krieg-Brückner, B., Shi, H. & Ross, R. (2004), 'A safe and robust approach to shared-control via dialogue', *Journal of Software* **15**(12), 1764 –1775.

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

Mann, W. C. (1983), An overview of the PENMAN text generation system, *in* 'Proceedings of the National Conference on Artificial Intelligence', AAAI, pp. 261–265. Also appears as USC/Information Sciences Institute, RR-83-114.

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

Mann, W. C. & Matthiessen, C. M. I. M. (1983), Nigel: A systemic grammar for text generation, Technical Report RR-83-105, USC/Information Sciences Institute. This paper also appears in a volume of the *Advances in Discourse Processes Series*, R. Freedle (ed.): *Systemic Perspectives on Discourse: Volume I.* published by Ablex.

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

Matthiessen, C. M. I. M. & Bateman, J. A. (1991), *Text generation and systemic-functional linguistics: experiences from English and Japanese*, Frances Pinter Publishers and St. Martin's Press, London and New York.

Niles, I. & Pease, A. (2001), Toward a standard upper ontology, *in* C. Welty & B. Smith, eds, 'Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001)', Association for Computing Machinery, Ogunquit, Maine.
URL: *http://projects.teknowledge.com/HPKB/Publications/FOIS.pdf*

Siegel, A. & White, S. (1975), The development of spatial representations of large-scale environments, *in* H. Reese, ed., 'Advances in child development and behavior', Academic Press, New York, pp. 9–55.

Tenbrink, T. (2005*a*), Identifying objects on the basis of spatial contrast: an empirical study, *in* C. Freksa, M. Knauff, B. Krieg-Brückner, B. Nebel & T. Barkowsky, eds, 'Spatial Cognition IV: Reasoning, Action, Interaction. International Conference Spatial Cognition 2004, Proceedings', Springer, Berlin, Heidelberg, pp. 124–146.

Tenbrink, T. (2005*b*), Semantics and application of spatial dimensional terms in English and German, Technical Report, SFB/TR8 Spatial Cognition No. 004-03/2005, University of Bremen.

Tenbrink, T. & Klippel, A. (2005), Achieving reference via contrast in route instructions and spatial object identification, *in* 'Workshop on Reference, 21st Scandinavian Conference of Linguistics, June 1-4, 2005'.