

A Minimal Recursion Semantic Analysis of Locatives

Fredrik Jørgensen*
University of Oslo

Jan Tore Lønning**
University of Oslo

The article describes a pilot implementation of a grammar containing different types of locative PPs. In particular, we investigate the distinction between static and directional locatives, and between different types of directional locatives. Locatives may act as modifiers as well as referring expressions depending on the syntactic context. We handle this with a single lexical entry. The implementation is of Norwegian locatives, but English locatives are both discussed and compared to Norwegian locatives. The semantic analysis is based on a proposal by Markus Kracht (2002), and we show how this analysis can be incorporated into Minimal Recursion Semantics (MRS) (Copestake et al. 2005). We discuss how the resulting system may be applied in a transfer-based machine translation system, and how we can map from a shallow MRS representation to a deeper semantic representation.

1. Introduction

Locative prepositional phrases (PPs) pose several puzzles, both to syntax and semantics. First, locatives may be either static (Example (1)), directional (Example (2)), or ambiguous (Example (3)):

- (1) Kim slept in Paris.
- (2) Kim is driving into Paris.
- (3) The mouse ran under the table.

As we see, the PP *under the table* may locate the whole event (Example (4)), it may express the goal of motion (Example (5)), or it may express (parts of) the path of motion (Example (6)):

- (4) The mouse ran around under the table.
- (5) The mouse ran under the table and stayed there.

* ILN, PO Box 1102 Blindern, 0317 Oslo, Norway. E-mail: fredrik.jorgensen@iln.uio.no.

** IFI, PO Box 1080 Blindern, 0316 Oslo, Norway. E-mail: jtl@ifi.uio.no.

Submission received: 29 June 2006; revised submission received: 16 November 2007; accepted for publication: 26 January 2008.

- (6) The mouse ran under the table into a hole in the wall.hole in the wall.

Second, we may use a combination of prepositions in order to express, for example, the source of motion, as seen in Example (7):

- (7) A mouse appeared from under the table.

Here, *under the table* seems to refer to a location, and it is *from* that expresses the directional component of the locative.

Third, we may see differences in the local interpretation of a PP and the same PP in a wider context. One popular example is the verb *put*, where both a static and a directional (goal) PP complement is accepted, and where both types of PP complements are interpreted as the goal of the motion.

- (8) Kim put the book on/onto the table.

Fourth, a preposition may not always be followed by an argument; consider *down* in Example (9). When it occurs without an NP complement it may be followed by another PP as in Example (10). How can the semantics of this construction be described compositionally?

- (9) A child ran down.

- (10) A child ran down under the bridge.

Another example of this phenomenon is how intransitive locatives disambiguate between different readings in Norwegian. The locatives *inne* and *inn* are static and directional (goal), in relation to a location which is *inside* in some (contextually determined) sense. When one of these locatives is succeeded by a locative PP, the second PP is given the same interpretation with respect to motion, as shown in Examples (11) and (12).

- (11) Musa løp inne_{static} i_{static} hullet.
 Mouse.DEF ran inside_{static} in_{static} hole.DEF
 'The mouse ran inside_{static} the hole.'

- (12) Musa løp inn_{goal} i_{goal} hullet.
 Mouse.DEF ran inside_{goal} in_{goal} hole.DEF
 'The mouse ran into the hole.'

Furthermore, new readings are available when a PP is preceded by *inn*, as seen in Examples (13) and (14).

- (13) Helikopteret fløy over_{static.or.path} byen.
 Helicopter.DEF flew over_{static.or.path} city.DEF.
 'The helicopter flew over_{static.or.path} the city.'

- (14) Helikopteret fløy inn over_{goal} byen.
 Helicopter.DEF flew in over_{goal} city.DEF.
 'The helicopter flew [to above]/[in over] the city.'

Finally, we see a variation in how different languages express locatives. Some languages have rich case systems, whereas others use adpositions to express locatives. The focus of this article is locative prepositions in Norwegian and English. We will not consider locative case systems here, but we will see differences in which locatives are lexicalized as prepositions, as shown in Example (15), where the Norwegian locative *derfra* is a lexicalization of the corresponding English complex PP *from there*.

- (15) Musa kom derfra.
 Mouse.DEF came there-from.
 ‘The mouse came from (over) there.’

Our goal in this article consists of three parts. First, we take recent developments within formal semantic approaches to locatives (in particular Kracht 2002) and show how these can be implemented in Minimal Recursion Semantics (MRS; Copestake et al. 2005), a flat and computationally tractable semantic meta-language for first-order logic. Second, we determine the degree to which the insights from this particular approach to locatives can be implemented in a non-transformational, unification-based computational grammar based on HPSG (Pollard and Sag 1994). And third, we investigate what consequences the current analysis has for syntax and the syntax–semantics interface. Our implementation is of a fragment of Norwegian. Norwegian is in many respects similar to English, and where they are similar we will use examples from English. There are also some interesting differences, and we will use them to illustrate how the new representations can be exploited in an experimental semantic transfer-based machine translation system, like the LOGON system (Oepen et al. 2004), which uses transfer representations based on MRS.

The rest of the article is organized as follows. In Section 2, we consider various approaches to locatives, before we describe the formal semantic approach proposed by Kracht (2002) and introduce MRS in Section 3. In Section 4, we start to present our own solution to how these two approaches can be combined. We proceed to classify Norwegian locative adverbs and prepositions and show how various constructions can be handled in Section 5. In Section 6, we describe the main results from our implementation of a grammar producing the intended Kracht-style MRS representations. In Section 7, we argue that the current analysis is useful for a range of NLP applications, and sketch how it can be applied in a machine translation system. Finally we compare our approach to other computational approaches in Section 8, discuss evaluation of the analysis in Section 9, and conclude in Section 10.

2. Syntax and Semantics of Locatives

In this section, we consider relevant linguistic literature on locatives which will serve as a background for our proposal. We will return to a comparison between our approach and other computational linguistic approaches in Section 8.

2.1 Relevant Issues

There is a growing literature on language and space, but not all of it is relevant for our purposes. We are interested in how locatives should be handled in a computational grammar with a compositional semantic component, and in particular how static and

directional locatives are treated. There are basically three types of questions relevant to our article, for which we will review relevant linguistic literature:

- Q1. **Formal Semantics: What do locatives denote?**
Are locatives referential (denoting a location) or modificational (denoting a property of events and entities)?
- Q2. **Syntax: Are locative PPs complements or adjuncts?**
If we want our grammar to have compositional semantics, this question is closely related to the previous one. Analyzing locative PPs as adjuncts is hard to combine with the referential interpretation of locatives, as this either requires that the modified phrase also denotes a location, or that some additional machinery is introduced to map the location to a property.
- Q3. **How is the distinction between static and directional locatives handled?**
Static and directional locatives differ both in syntactic distribution and truth conditions, and as we have observed, occurrences of locatives may be ambiguous between the two.

A prominent question in the literature on locative prepositions is the **geometric structure** of particular prepositions: for example, what geometric relationship between the ball and the cup has to be realized for an expression like *the ball is in the cup* to be true (e.g., Herskovits 1986)? As interesting as these questions are, we think they belong to the lexicon and the lexical semantic domain and fall outside the scope of this article. There is also interesting work which relates *the meaning of individual prepositions* to various types of psychological studies; in particular, studies of under which conditions the various prepositions are used (see, e.g., many of the references in Bloom et al. [1996] and in van der Zee and Slack [2003]). But again this plays a complementary role to our work.

One particular question in the lexical semantics of prepositions is the distinction between what Zwarts and Winter (2000), following Herskovits (1986), call **projective** and **non-projective** prepositions. A non-projective preposition, like *outside*, requires only spatial knowledge of the location of the two objects, while a projective preposition, like *behind*, requires some further information about directions from the reference object. Levinson (1996) makes a further distinction between three types of reference frames called (i) **intrinsic**, where *behind the house* means on the other side of the house than what would be classified as *the front side of the house*, (ii) **extrinsic**, where *behind the house* is on the opposite side of the house than the speaker, and (iii) **absolute**, which applies to expressions like *north of the house*. We also think that these distinctions belong to the lexical semantic domain.

2.2 Hjelmslev's Theory of Cases

One of the earliest contributions to the study of locatives is Hjelmslev's (1935) theory of cases. Hjelmslev views case as a relation between two objects (Q1 above), which may be nominal or verbal (Q1/Q2). Cases may be either complements or modifiers (Q2), and are expressed through adpositions or inflections. These observations are based on data from a wide range of languages, including languages with locative case systems, such as Hungarian and the Caucasian language Tabasaran. Hjelmslev proposes a three-dimensional case system (Q3), where the dimensions express **directionality**, **coherence**, and **subjectivity**. The directional dimension consists of **approach**, **neutrality**, and **sep-**

aration. The coherence dimension has two interpretations, **inherence** (enclosure) and **adherence** (contact). And finally, the subjectivity dimension describes the dependence relation between the point of view and the location, being either **subjective**, **objective**, or **neutral**. For example, whereas *behind* or *left of* is dependent on the point of view, *under* is not, according to Hjelmlev.

Hjelmlev's theory of cases relates to all three questions, and it has formed the basis of much of the later work done on locatives. Even though we will focus on languages where locatives are expressed by prepositions in adpositions, and ignore languages where locatives are expressed through inflectional case, it is a point for us to base our approach on formalisms that also extend to languages where locatives are expressed by case. When it comes to the dimensions, directionality is particularly important for the questions we raised in the introduction, while the latter two dimensions relate to lexical semantics, which we will not pursue here.

2.3 Bierwisch's Grammar of Locatives

Bierwisch (1988) gives a detailed account of both syntactic and semantic aspects of locative prepositions in German, and is particularly interested in the relationship between the denotation of locatives (*Q1*) and the consequences for the syntactic treatment of locatives (*Q2*). Bierwisch notes that locative PPs serve as **predicates** (Example (16)), **arguments** (Example (17)), and **modification adjuncts** (Examples (18)–(19)) (examples taken from [1988, page 5]; our translations).

- (16) Er ist in der Schule.
He is in school.
- (17) Der Brief liegt auf dem Tisch.
The letter is lying on the table.
- (18) Ich kaufe das Buch in Berlin.
I will buy the book in Berlin.
- (19) eine Brücke über die Moldau
a bridge over Moldau

Bierwisch notes that even though predication, complementation, and modification are clearly distinct syntactic relations, distinguishing between locatives as adverbial or optional complements is in many cases arbitrary (*Q2*). Bierwisch argues that directionality is a syntactic feature, $\pm\text{DIR}$, on locatives. But as we shall see, a more complex inventory of syntactic features corresponding to different types of directionality is needed.

Bierwisch (1988) discusses two different positions with respect to the denotations of locative PPs (*Q1*): The **referential interpretation**, where locative PPs denote regions, just like NPs denote things, and the **modification interpretation**, where the locative PPs denote properties of being located at a certain place:

The referential interpretation seems to be appropriate for PPs in argument position, as e.g., *Hans liegt im Bett* can plausibly be said to express a relation between *Hans* and a place denoted by *im Bett*. It is difficult to see, however, how on this account PPs can serve as modifiers or predicatives — unless a place is construed as a property, but that would violate the gist of the referential interpretation. The modification interpretation, on the other hand, concerns itself with PPs as adjuncts, but seems to be

in trouble with PPs in argument position. From this, one might be tempted to draw the conclusion that both interpretations are partially right and that they both are needed. (Bierwisch 1988, page 8)

We will not discuss predicatives in this article, but refer to Kracht (2002) for treatment of adnominal locatives as a property of individuals. The duality with respect to the interpretation of locatives as referential or modificational is central to our article, and we will show, building on Kracht, how this can be handled in a computational grammar.

2.4 Jackendoff’s Conceptual Semantics

Ray Jackendoff’s (1983, 1990) **conceptual semantics** is a decompositional theory of meaning, heavily influenced by X-bar theory. Conceptual semantics organizes a repertoire of major conceptual categories, the **semantic parts of speech**, into function–argument structures. Lexical entries are encoded as Lexical Conceptual Structures (LCSs), as exemplified by the entry for the preposition *in* in (Example (20)), where we find the lexeme, the part-of-speech category, the selectional restriction, and the semantics on four separate lines.

Particularly interesting in our context is the semantic analysis of static and directional locatives (*Q1/Q3*), and how they combine with verbs (*Q2*). The LCSs of static locatives are claimed to have one layer, a Place function (Example (20)), as opposed to the LCS of directional locatives, which have two layers, a Path function as the outer function, and a Place function as the inner function (Example (21)). According to Jackendoff, there are five different Path functions: TO, FROM, TOWARD, AWAY-FROM, and VIA. These map from a Place to a Path.

$$(20) \left[\begin{array}{l} \text{in} \\ \text{P} \\ _ \text{NP}_j \\ [\text{Place IN} ([\text{Thing }]_j)] \end{array} \right]$$

$$(21) \left[\begin{array}{l} \text{into} \\ \text{P} \\ _ \text{NP}_j \\ [\text{Path TO} ([\text{Place IN} ([\text{Thing }]_j)])] \end{array} \right]$$

Furthermore, Jackendoff (1990, page 45) treats directional locatives as arguments to motion verbs, such that, for example, *run* subcategorizes for an optional directional locative argument (*Q2*), where the optionality is represented with angle brackets, note the LCS for *run* given in Example (22). In the analysis of Example (23) the optional PP argument of *run* is coindexed with the PP *into the room*. Through co-indexing (*j*) this results in the conceptual structure in Example (24), where the path denoted by the PP enters into the function $\text{GO}(\text{subject entity}, \text{path})$ and thereby expresses the traversing of the path. By convention, the entity indexed *i* is taken to be the subject position.

$$(22) \left[\begin{array}{l} \text{run} \\ \text{V} \\ _ \langle \text{PP}_j \rangle \\ [\text{Event GO} ([\text{Thing }]_i, [\text{Path }]_j)] \end{array} \right]$$

(23) [_S [_{NP} John [_{VP} ran [_{PP} into [_{NP} the room]]]]

(24) [_{Event} GO ([_{Thing} JOHN], [_{Path} TO ([_{Place} IN ([_{Thing} ROOM]))])]]

Jackendoff's description clearly shows that matters are more complex than the usual static/directional distinction. First, he introduces a finer classification between the different directionals. Second, he proposes there be a common core, the place, to different prepositions (e.g., *in* and *into*) and to the different uses of a preposition (e.g., the place-use and the TO- and VIA-uses of *under*). Both observations will be part of our analysis.

But there are several problems with Jackendoff's approach. First, one will have to specify for any verb which can be modified by a directional PP, that the modifier is an optional part of the verb's semantics (or LCS). Because one and the same verb (e.g., *run*) can be combined with several directional PPs at the same time, as in Example (6), one would need a way to combine the semantics of the different locative PPs. Whether this should be done by merging the directionals into a single path, or by accepting multiple PP complements, is unclear in Jackendoff's theory. We believe incorporating the optional modifier's semantics into the verb's lexical entry leads to a very elaborate and inflexible semantics, and would prefer to avoid this, if possible.

The second, and more serious, weakness is that static and directional locatives are treated differently in the syntax. Static locatives are regarded as adjuncts, whereas directional locatives are regarded as complements of motion verbs. Unless there is provided syntactic evidence for this distinction, this violates the autonomy of syntax, as the semantics of the construction (whether it is the event or its participants being located) governs the syntactic analysis (whether the construction is analyzed as an adjunct or a complement). There are of course different syntactic restrictions on what kind of verbs static and directional PPs can modify, but this applies to many modifying PPs, and is not in itself sufficient to warrant different syntactic analyses. We believe that a uniform treatment of locatives, giving the same syntactic analysis independent of directionality, would provide a more sound and consistent analysis.

2.5 Fong's Directionals as Diphasic Events

Fong (1997) analyzes directional locatives. She focuses on the locative cases of Finnish, but includes comparisons to other languages, including English. Her main focus of interest is on how directional locative case in Finnish and directional prepositions in English can be used when there is no movement involved. In Finnish one uses expressions which literally when translated into English would become *forgot my wallet into my car*, and in English we have formulations like *a bridge from Buda to Pest*. Whereas other approaches take the spatial use as basic, and consider the more abstract uses as derived, Fong takes the opposite stand. The locative case is not in itself about paths. It has a more abstract meaning, and can be used when there is an underlying ordered domain, together with a proposition changing truth value, either from false to true or from true to false along the ordering. This abstract structure can be mapped to a more concrete structure, including space in English, and space or time in Finnish. Thus in Example (25) there is an ordering in space such that Example (26) is false at the beginning of this ordering, and true towards the end of the ordering.

(25) He went into the room.

(26) He is in the room.

Kracht (2002) compares his approach to Fong's. One problem is that Fong's approach does not extend to the VIA-reading (in Jackendoff's terminology) of sentences like Example (3). From our point of view, it is also a point that she does not describe a compositional analysis. Also our focus in this article is strictly on the spatial readings of the locatives.

2.6 Zwarts's Vector Space Semantics

In a series of papers, Zwarts (1997, 2003a, 2003b, 2005; Zwarts and Winter 2000) has developed a semantics for locatives based on vectors. A main motivation for this proposal is the observation that locatives may be modified by phrases expressing distance or direction; for example, a measure phrase (*ten meters*), an adverb (*diagonally*), or a dimensional adjective (*far*), as seen in Example (27).

(27) Superman is far/diagonally/ten meters above the building.

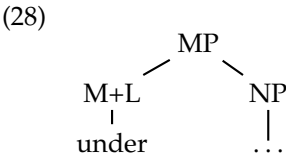
Many approaches to locatives will consider *above the building* to denote a region and let this be a set of points. It is then not obvious how the modifier *ten meters* can apply to this region. Zwarts and Winter (2000) propose instead to let the locative PP denote a **set of vectors** (Q1). Vectors are directed line segments between points in space. Loosely speaking, we may say that *above the building* denotes the set of vertical vectors starting in the building and ending above it. By then letting *ten meters* denote the vectors having a length of ten meters, the modification can be considered set intersection, and the result are all vectors of length ten meters starting in the building and ending vertically above it. Syntactically, Zwarts and Winter consider locatives to be modifying adjuncts. They mainly consider how nouns are modified. Because nouns denote sets of individuals, and locatives denote sets of vectors, they assume a primitive semantic function *loc* which maps individuals to their regions as sets of points (or vectors). This is used for the landmark (e.g., *the building* in the locative PP *above the building*). When a locative modifies a noun, the inverse function is used, such that the set of vectors is turned into the set of objects which are located within the set of (endpoints of) the vectors.

Even though this approach seems to solve the semantic problem of this particular type of construction, we are not sure how general it can be made. The compositional analysis of Zwarts and Winter (2000) is applied to modification of nouns and to predicative uses of locatives; they do not consider verbal modification. We can see how the approach can be adopted to verbal modification of static locatives in an event based framework by introducing a function *loc_e*, mapping events to regions, but we do not see how it extends to directional locatives. In later articles, Zwarts (2003b, 2005) extend the approach to the directional PPs. A directional PP does not denote a set of vectors, but a set of paths. A path can be considered a sequence of vectors. For example, in *John went to the store*, the path John followed can be formalized as a sequence of vectors, $\{v_i | i \in I\}$ for an ordered set I , where each v_i has the store as its starting point, where John successfully traverses their endpoints, and where the last vector also has the store as its endpoint. Intuitively this is a correct description, but Zwarts (2003b, 2005) does not show how this can be handled in a compositional grammar. It is a particular problem for a compositional treatment that static PPs and directional PPs are ascribed denotations of different types, unless they are also given different syntactic analyses (cf. discussion on Jackendoff in Section 2.4).

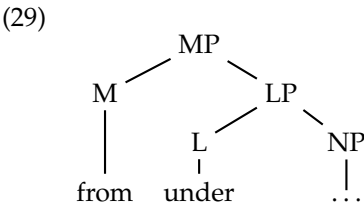
2.7 Kracht’s Two-Layered Locatives

In Kracht (2002), a novel approach to the semantics of locatives is described. Kracht presents a solution to Bierwisch’s (1988) problem of the duality of locatives, by claiming that locatives consist of two layers (Q1/Q2). The lower layer, called a **localizer phrase** (LP), defines a region or location, while the higher layer, called a **modalizer phrase** (MP), defines events of motion with respect to this region (Q1).

Both from a semantic and syntactic point of view, Kracht claims a locative expression is structured as [MP M [LP L NP]]. Here, L is a localizer which together with the NP forms an LP whose semantic contribution is a location. M is a modalizer which together with the LP forms an MP, whose semantic contribution is an event modifier. The analysis is motivated on the one hand by the dualistic nature of locative semantics, and on the other hand by syntactic and morphological data. Locatives are realized differently in different languages. Some languages use pre- or postpositions, whereas other languages use locative case. In particular, case languages like Finnish, Hungarian, and Tsez realize the two different layers as different suffixes. Still, whether the locative is expressed by an adposition or case, the M+L most often form a unit. This is by far the most frequent case for English, meaning that the preposition *under* in Example (3) is analyzed as shown in Example (28).



In English, the two-layering is usually not explicit, but it is found in constructions like Example (7), corresponding to:



MPs denote sets of events, and their syntactic function is that of adverbial modification. The data Kracht discusses give support for five different types of modalizers or modes, which are the heads of the MPs, denoting five different types of events.¹ Thus, if Kracht is right, the modes are true semantical interlingua predicates. The modes are:

1. **static (st)**, corresponding to Jackendoff’s Place Function
2. **coinitial (ci)**, corresponding to Jackendoff’s Path Function FROM
3. **cofinal (cf)**, corresponding to Jackendoff’s Path Function TO

¹ Kracht (2002) also mentions a sixth mode, the recessive mode, corresponding to Jackendoff’s AWAY-FROM. But the data he discusses do not sufficiently support the existence of this mode.

4. **transitory (tr)**, corresponding to Jackendoff's Path Function VIA
5. **approximative (ap)**, corresponding to Jackendoff's Path Function TOWARDS

Kracht also claims that LPs only occur in restricted contexts, for example, as a complement selected by a verb or a preposition. The LP denotes a location, and the head of an LP is a localizer function, which defines a location in relation to its complement.

Furthermore, Kracht claims that verbs may select a PP complement of a certain mode syntactically, but only interact with the LP semantically. In other words, the verb takes a location as the semantic argument. This is explained in more detail by Kracht (2006). The principle governing this particular verb–PP interaction is called the **Emptiness Principle** (where X corresponds to a preposition, and C to a PP for our purposes):

Emptiness Principle. Suppose that X is a syntactic marker in the constituent C . Suppose further that the presence of X in \bar{C} is determined purely by non-semantic rules (for example selection, agreement, Sandhi). Then the meaning of X is empty, namely the identity function. (Kracht 2002, page 204)

Thus, we end up with a theory that explains both the dualistic nature of prepositions semantically, and how the two-layering interacts with the syntax.

2.8 Summary

In this section, we have looked at several theories on locatives. These theories each have a different focus and different goals. Returning to our initial questions, we see that both Kracht (2002) and Jackendoff (1990) preserve the two-layered nature of locatives. But where Kracht's solution allow directional locatives to be modificational adjuncts, Jackendoff is committed to treating directionals as referential complements of motion verbs. Furthermore, Fong's (1997) diphasic events are insufficient to capture the nature of Kracht's transitory and approximative modes. In addition, neither Jackendoff nor Fong provides a formalized compositional semantics to the same extent as Kracht. Furthermore, we find Kracht's approach to be of greater cross-linguistic validity, and porting the analysis to languages with locative case or a mix of case and prepositions is therefore possible. For these reasons, we have chosen Kracht's proposal as the basis for our implementation of locatives.

3. Formal Background

In this section, we describe the formal semantic aspects of Kracht's (2002) proposal, as well as the framework of MRS (Copestake et al. 2005), a meta-level language for logical representations.

3.1 Formal Ontology and Formal Semantics

Kracht (2002) argues for his analysis in terms of modalizers and localizers from data—in particular data regarding case—in several typologically different languages. He proposes a formal ontology in which he specifies a formal semantics for the localizers and modalizers, and furthermore shows how the fact that a verb can select an LP or NP can be implemented in a grammar. Along the way he discusses several other phenomena,

like the lexical meaning of the different localizers, with varying degrees of rigor. We will not have much to say about case or languages other than English and Norwegian. What we will do here is to present Kracht's formal ontology and semantics for the locatives which will serve as a starting point for our analysis. When it comes to the integration in a grammar, we will depart from his proposal. One should also keep in mind that his ambition is not to give full-fledged analyses of sentences containing locatives; he only discusses the relative contribution from the locatives.

To describe the semantics formally, Kracht uses a typed lambda calculus with the following basic types:

- t, the type of truth values
- e, the type of objects
- v, the type of events
- i, the type of time points
- p, the type of spatial points
- r, the type of regions
- j, the type of intervals

There are two ways to construct complex types. If α and β are types then so are $\alpha \rightarrow \beta$, and $\alpha \times \beta$. We write $\phi : \tau$ to express that the term ϕ is of type τ . A term of type t is a formula. Its denotation in a model will be a member of the domain D_t , the set of truth values. Each model will contain a non-empty set of individuals D_e and each term of type e denotes a member of this domain. A simple formula, for example, *man(cain)* of type t can be constructed if *man* : (e \rightarrow t) and *cain* : e. The \times -constructor allows a predicate to take more than one argument, for example, *brother* : (e \times e) \rightarrow t, and *brother(cain, abel)* : t. The domain of $\alpha \rightarrow \beta$, $D_{(\alpha \rightarrow \beta)}$, is the set of all functions from D_α to D_β . The domain of $\alpha \times \beta$, $D_{(\alpha \times \beta)}$, is $D_\alpha \times D_\beta$. Functions are constructed by the use of lambda in the normal way: if $\phi : \alpha$, x is a variable, and $x : \beta$ then $\lambda x \phi : \alpha \rightarrow \beta$.

The events allow for a Davidsonian treatment of verbs and adverbs, where a sentence like Example (30) denotes a set of events as in Example (31).

(30) Kim is driving.

(31) $\lambda e[\text{drive}(e, \text{kim})]$

The domain of events is a non-empty set D_v . Kracht assumes the set of time points to be (isomorphic to) the set of reals, $D_i = \mathbb{R}$, and the set of space points to be (isomorphic to) the Cartesian space $D_p = \mathbb{R}^3$. The type j is assumed to be a sub-type of $i \rightarrow t$, and its domain, $D_j \subseteq D_{(i \rightarrow t)}$, is the set of intervals. The type of regions, r, is assumed to be a sub-type of $p \rightarrow t$. Its domain $D_r \subseteq D_{(p \rightarrow t)}$ is the set of all path-connected sets of points. For our purposes, we do not have to go further into the details here.²

2 We have made one simplification compared to the original proposal. We have not included groups for modeling plurals because it is not relevant for the points we want to make in this article. Also, our notation departs in some respects from the original one. Kracht uses partly curly bracket notation for (characteristic functions of) sets, whereas we stick to a lambda notation.

For adverbial modification, Kracht claims the MP in Example (32) denotes a set of events, as shown in Example (33). The nuclear sentence in Example (30) also denotes the set of events in Example (31) and the modification works as an intersection between the two sets. Hence Example (34) is represented as Example (35), where we use double quotes to mark material which we have not yet fully analyzed.

(32) into Paris.

(33) $\lambda e[\text{"into Paris"}(e)]$

(34) Kim is driving into Paris.

(35) $\lambda e[\text{drive}(e, \text{kim}) \wedge \text{"into Paris"}(e)]$

We now turn to the inner structure of the MP and the semantics of the LP. A sentence like Example (34) says something about the relationship between the region where Kim is driving and the region within the city of Paris. To achieve this, Kracht assumes a primitive function taking objects to their regions, *LOC*. As objects can be moving, this function has to be time-dependent, hence of type $e \rightarrow (i \rightarrow r)$. Thus if t is a particular time point, $t : i$, and the term *paris* of type e is the name of Paris, *LOC paris*(t), which is of type r , expresses the region of Paris at time t .

The LP cannot denote such a simple region, however, but has to be a set of regions, what Kracht calls a **neighborhood**. In the example, this will be the set of all regions that can be claimed to be in Paris.³ Moreover, such neighborhoods may also vary with time, cf. *behind the car*. Hence the denotation of the LP has to be a parameterized neighborhood, of type $i \rightarrow (r \rightarrow t)$.

An MP has the structure [M [L NP]]. If we assume the contribution from the NP to be of type e , then the localizer, *L*, will be of type $e \rightarrow (i \rightarrow (r \rightarrow t))$, and the modalizer, *M*, of type $(i \rightarrow (r \rightarrow t)) \rightarrow (v \rightarrow t)$.

To take an example, the English prepositions *into* and *in* have different modes—*into* is cofinal whereas *in* is ambiguous between static and cofinal mode—but they share the same localizer function *IN*. The common LP corresponding to the expressions *in Paris* and *into Paris* will be *IN paris*. We could consider such localizers as primitive, but Kracht takes one step further and assumes that there are basic relations between regions, called local relations. For example, i , of type $(r \times r) \rightarrow t$, is a particular relation between two regions locating the second region within the first region. The localizer *IN* can then be analyzed as $\lambda x \lambda t \lambda r [i(\text{LOC}(x)(t), r)]$ and the LP corresponding to *IN paris* as $\lambda t \lambda r [i(\text{LOC}(paris)(t), r)]$, where r is a variable ranging over regions. Other prepositions contribute a different local content. Kracht thinks there is no upper bound on the possible different localizers in a language.

If we write *L* for the logical representation of the LP, the static MP can be written *ST(L)*. Here *ST* is a modalizer, a function of type $(i \rightarrow (r \rightarrow t)) \rightarrow (v \rightarrow t)$. Kracht argues from cross-linguistic data for a finite set of modalizers, in addition to the static mode, these being the four modes for directional PPs: coinitial, cofinal, transitory,

3 In this particular example, it could have worked to let *in Paris* denote the maximal region which is in Paris, and to say that other regions were in Paris if they were subregions of this region. But that approach does not work for *The book is on the table*. For the book to be on the table, it does not suffice that the location of the book is contained in the maximal region which is on the table. The region where the book is located has to make contact with, and be supported by, the table. Moreover, the sentence might be true even though the whole book is not within the maximal region on the table.

approximative. They each contribute a different function CI , CF , TR , AP (corresponding to M) of the same type as ST . These modalizers are shorthand for more structured functions which can—at least partly—be expressed in the lambda calculus. To achieve this, Kracht assumes some basic functions. In addition to $LOC : (e \rightarrow (i \rightarrow r))$ introduced previously, there is a function $TIME : (v \rightarrow (i \rightarrow t))$, taking an event to the set of time points in that event, a function $\mu : (v \rightarrow e)$, yielding the principal mover of an event, and a function $LOC_v : (v \rightarrow r)$, yielding the (time independent) region of an event.⁴ In addition, there are several defined functions used, for example, CF^* , Π , $PEND$, to make the representations simpler, and hiding some of the complexity. The cofinal reading $CF(L_1)$ can then be spelled out as a series of equivalent representations (Examples (36)–(42)) making use of such defined functions and relations.⁵

$$(36) \text{ CF}(L_1)$$

$$(37) \lambda L \lambda e [CF^*(\mu(e), L, TIME(e))](L_1)$$

$$(38) \lambda e [CF^*(\mu(e), L_1, TIME(e))]$$

$$(39) \lambda e [PEND(\Pi(\mu(e), L_1) \cap TIME(e), TIME(e)))]$$

$$(40) \lambda e [PEND(\lambda t [L_1(t)(LOC(\mu(e)(t))], TIME(e)))]$$

$$(41) \lambda e [\neg \forall t (TIME(e)(t) \leftrightarrow L_1(t)(LOC(\mu(e)(t)))) \wedge \\ \forall t (TIME(e)(t) \rightarrow \exists t' (TIME(e)(t') \wedge t \leq t' \wedge L_1(t')(LOC(\mu(e)(t')))))]$$

$$(42) \lambda e [\exists t \in TIME(e) (\neg L_1(t)(LOC(\mu(e)(t)))) \wedge \\ \forall t (TIME(e)(t) \rightarrow \exists t' (TIME(e)(t') \wedge t \leq t' \wedge L_1(t')(LOC(\mu(e)(t')))))]$$

In plain English, this says that the mover of the event starts in a region which is not one of the regions in the neighborhood denoted by L , and ends in a region which is in the neighborhood denoted by L . Similar definitions can be given for the coinital reading and for the transitory reading. To define the approximative reading, the representation language has to be extended with a function representing distance in the underlying metrical space. The static reading, $ST(L)$, can through a series of equivalences be spelled out as the following:

$$(43) \text{ ST}(L) = \lambda e [\forall t (TIME(e)(t) \rightarrow (L(t)(LOC_v(e)))]$$

For any time point t during the event e , the location of e belongs to L .

3.1.1 Comparing Kracht's and Zwarts' Proposals. In Section 2, we compared Kracht's proposal briefly with some of the earlier proposals from the literature. Having the formal apparatus on board, we will now compare it in more detail to Zwarts and Winter's (2000) proposal, to investigate if and how the two approaches can be combined. One possibility is to try to merge the two proposals. We may extend Kracht's semantics with

⁴ Kracht uses the same function for LOC and LOC_v , but we have simplified a little as it does not matter for our purposes.

⁵ We refer to Kracht (2002) for the details of the calculations and functions in Examples (36)–(42).

vectors and let p be the type of vectors instead of points. The domain of regions, D_r , would still be a subset of $D_{(p \rightarrow t)}$. A set of vectors would be a region if the set of endpoints of the vectors is a path-connected set of points, and so forth.

Zwarts and Winter's treatment of modification as intersection, as in *ten meters above the house*, can be reconstructed within this framework if we consider it as LP-modification. This means that *ten meters*, which intuitively is a set of vectors, V , hence of type $p \rightarrow t$, has to be lifted to type $i \rightarrow (r \rightarrow t)$. One can take the temporal part to be constant. For the regional part, a possibility is to admit all regions consisting of vectors which are ten meters long, hence lift V to the following:

$$(44) \lambda t \lambda r [\forall v (r(v) \rightarrow V(v))]$$

This is a little simplified, though, because for a hot air balloon to be ten meters above the house one cannot request all parts of the balloon to be ten meters above the house. But that is as much a challenge for Zwarts' original proposal and it does not belong to the compositional semantics to solve.

This analysis has a problem with compositionality, though. In an expression like *under the top*, the preposition *under* is interpreted as M+L, and *under the top* as an MP. We do not have a constituent corresponding to the LP part of *under the top*, which can be modified by *ten meters*. Zwarts and Winter's solution works, on the other hand, because they only consider static PPs, or rather NP-modifying PPs. It remains to show how this can be extended to directional PPs. We will hence ignore such modifiers for the rest of this article.

3.2 Introducing Minimal Recursion Semantics

MRS (Copestake et al. 2005) is a framework for semantic representations in large-scale grammars. The main goals for MRS is to be expressively adequate allowing linguistic meaning to be represented properly, to be compatible with other modules in a grammar, to be computational tractable, and to support underspecification. The semantic representations in MRS can be considered a meta-level language for representing an underlying object language which is typically first-order logic with generalized quantifiers. Example (46) is an MRS structure for the sentence in Example (45). The basic building blocks are the elementary predications (EPs), like *dog(x)* and *chase(x,y)*, corresponding to atomic formulas.

(45) Every dog chases some cat.

$$(46) \langle h0, h7, \{h1: \text{every}(x, h2, hA), h3: \text{dog}(x), h4: \text{some}(y, h5, hB), h6: \text{cat}(y), h7: \text{chase}(x, y)\}, \{h0 =_q h7, h2 =_q h3, h5 =_q h6\} \rangle$$

MRS is often described as a flat semantics, because the MRS structures are syntactically non-recursive. This is achieved by **labeling** all EPs with **handles**. In Example (46), $h3$ is the label on the predication *dog(x)*. In addition, an EP may have handles as arguments (cf. the argument $h2$ of *every(x, h2, hA)*). Quantifiers introduce special relations in an MRS, corresponding to generalized quantifiers. The first argument of a quantifier relation is the bound variable, the second is the restriction, and the third is the body. These handles are employed in the mapping from the flat meta-level representation to the object-language expressions with explicit representations of scope.

Scope underspecification is represented in MRS by having handles in argument positions of EPs that are not the labels of other EPs. The handles are linked (or equated) in the construction of object-language expressions. The unconstrained handle arguments hA and hB in Example (46) may be linked in two ways to construct the expressions in Examples (47) and (49), corresponding to the FOL expressions in Examples (48) and (50), respectively.

$$(47) \text{ some}(y, \text{ cat}(y), \text{ every}(x, \text{ dog}(x), \text{ chase}(x, y)))$$

$$(48) \exists y(\text{ cat}(y) \wedge \forall x(\text{ dog}(x) \rightarrow \text{ chase}(x, y)))$$

$$(49) \text{ every}(x, \text{ dog}(x), \text{ some}(y, \text{ cat}(y), \text{ chase}(x, y)))$$

$$(50) \forall x(\text{ dog}(x) \rightarrow \exists y(\text{ cat}(y) \wedge \text{ chase}(x, y)))$$

To constrain the possible linkings of handles and thereby the possible scope relations, MRSs may contain so-called equality modulo quantifiers. The constraint for the handle h and the label l (of some EP), $h =_q l$, called a qeq constraint, expresses that either $h = l$, or some non-repeating chain of one or more labels of quantifier EPs intervene. An object-language formula is described by the MRS structure only if it respects all the constraints and all variables are bound.

The general form of an MRS can be considered a four-tuple $\langle GT, LT, R, C \rangle$ where R (relations) is the bag of EPs, and C (constraints) is a bag of all the qeq constraints. In addition, GT (global top) is the label of the topmost EP of the structure, while LT (local top) is the label of the topmost EP of a sub-structure. The LT is needed for rules of semantic composition. Furthermore, the GT is always qeq to the LT of the root phrase.

Semantic composition in MRS is performed as follows: The bag of EPs (R) and the bag of qeq constraints (C) of the daughter nodes are appended on the mother node. Furthermore, the global tops (GT) of the daughters are equated on the mother node. Scopal and intersective combination are treated differently. For intersective combination, the local tops (LT) are equated. For example, if *black* corresponds to the expression in Example (51) and *cat* corresponds to Example (52), then *black cat* corresponds to Example (53). (The coindexation of x and y is done through the interaction with the grammar.)

$$(51) \langle h0, h1, \{h1: \text{ black}(y)\}, \{\} \rangle$$

$$(52) \langle h0, h5, \{h5: \text{ cat}(x)\}, \{\} \rangle$$

$$(53) \langle h0, h1, \{h1: \text{ black}(y), h1: \text{ cat}(y)\}, \{\} \rangle$$

For scopal combination, the LT of the mother is equal to the LT of the scoping daughter, and a qeq constraint is added between the scopal argument of the scoping daughter and the label of the scoped-over daughter. One result is seen in Example (54), corresponding to *probably sleeps*.

$$(54) \langle h0, h1, \{h1: \text{ probably}(h2), h3: \text{ sleep}(x)\}, \{h2 =_q h3\} \rangle$$

For quantifiers, the scopal argument is the restriction of the quantifier, whereas the body of the quantifier is left unconstrained. This means that the handle variable in restriction position is qeq to the label of the restriction's LT . For example, if *every*

corresponds to the expression in Example (55) and *cat* corresponds to Example (56), then *every cat* corresponds to Example (57).

$$(55) \langle h0, h1, \{h2: \text{every}(x, h3, h4)\}, \{\} \rangle$$

$$(56) \langle h0, h5, \{h5: \text{cat}(x)\}, \{\} \rangle$$

$$(57) \langle h0, h1, \{h2: \text{every}(x, h3, h4), h5: \text{cat}(x)\}, \{h3 =_q h5\} \rangle$$

4. MRS Analysis of Locatives

Kracht's analysis is based on higher order functions and functional application. How can that be ported to a system based on (first-order) predication where composition is built by unification? And how shall the phenomena be described in a grammar? We will work our way top down starting with how MPs are combined with other phrases and then proceed to the inner structure of MPs. We will take as our starting point the large-scale computational English Resource Grammar, ERG (Flickinger 2002),⁶ an HPSG grammar using MRS as the format for semantic representation (see Flickinger 2002 and Section 6 of Copestake et al. 2005). ERG uses a Davidsonian approach where verbs introduce events, and adverbial intersective modification involves coindexation of the event variables of the verbal relation and the modifier relation. Because Kracht's approach is also event-based and he considers all adverbial locative PPs (i.e., MPs) as intersective modifiers, the analysis of MPs can be expressed straightforwardly in MRSs and is, in fact, similar to the analysis in ERG.

4.1 MRS Analysis of Modalizer Phrases (MPs)

The MP can be represented with an EP containing an event variable. Schematically, the MRS structures in Examples (58) and (59) are composed into Example (60). As we described in the previous section, intersective adverbial modification in MRS is expressed by coindexation of the labels of the intersected EPs. Shared labels are treated in MRS as implicit conjunctions, and coindexation of labels will thus have the same effect as conjoining the two EPs. (The event variables are coindexed through the interaction with the grammar.)

$$(58) [\text{A mouse appeared}]' = \\ \langle h0, h5, \{h1: \text{some}(x, h2, h3), h4: \text{mouse}(x), h5: \text{appear}(e, x)\}, \\ \{h0 =_q h5, h2 =_q h4\} \rangle$$

$$(59) [\text{from under the table...}]' = \\ \langle h6, h7, \{h7: \text{"from under the table"}(e')\}, \{\dots\} \rangle$$

$$(60) [\text{A mouse appeared from under the table}]' = \\ \langle h0, h5, \{h1: \text{some}(x, h2, h3), h4: \text{mouse}(x), h5: \text{appear}(e, x), \\ h5: \text{"from under the table"}(e)\}, \{h2 =_q h4, \dots\} \rangle$$

⁶ For more information on ERG, see <http://www.delph-in.net/erg/>.

4.2 MRS Analysis of Localizer Phrases (LPs)

We now turn to the question of the inner structure of the MP, the representation of the LP, and how the MP is constructed from the LP. This is more delicate, as in Kracht's formalism an L representing an LP is a higher order function—of type $(i \rightarrow (r \rightarrow t))$. This function occurs on the one hand as an argument to other functions, for example, to the modalizer functions, ST , CF , CI , TR , AP as in Example (36), and on the other hand as a functor (or relation) when the deeper semantics of the modalizers are spelled out, for instance, as in Example (42).

Our solution is based on a division of semantic representations into several layers. We will distinguish between a surface level and deeper levels. The surface level will represent all semantic information that is grammaticalized, and it will be appropriate for parsing and generation, and possibly other types of applications. The surface representation does not have to represent all the deeper semantic meaning explicitly, but the deeper meaning should be derivable from the surface representation. Thus, for this particular example, it suffices with a surface representation corresponding to Example (36) since we have demonstrated how Example (42) can be derived deterministically. Because an LP will always be argument to an M or to a verb, it is possible to give representations where we only see such L s in argument positions. We may therefore consider the type $(i \rightarrow (r \rightarrow t))$ of parameterized neighborhoods a primitive type l in our surface representation language. A modalizer function, like CI , takes first a parameterized neighborhood, L , and then an event, and yields a truth value. It is of type $(i \rightarrow (r \rightarrow t)) \rightarrow (v \rightarrow t)$. This corresponds to being of type $((i \rightarrow (r \rightarrow t)) \times v) \rightarrow t$, or $(l \times v) \rightarrow t$. Hence it can be represented in MRS as a relation, ci_mod , taking two arguments. Ignoring the analysis of *the table*, this yields the MRS representation in Example (62) for the MP in Example (61). The full semantic representation of Example (7) is shown in Example (63).

(61) $[_{MP}$ from $[_{LP}$ under $[_{NP}$ the table]]]

(62) $\langle h0, h1, \{h1: ci_mod(e, l), h1: under_loc(l, x), \}$

(63) $[A\ mouse\ appeared\ from\ under\ the\ table]' =$
 $\langle h0, h5, \{h1: some(x, h2, h3), h4: mouse(x),$
 $h5: appear(e, x), h5: ci_mod(e, l), h5: under_loc(l, y),$
 $h9: def(y, h6, h7), h8: table(y)\}, \{h0 =_q h5, h2 =_q h4, h6 =_q h8\}$

For applications, we can convert this simple surface interpretation to Kracht's deep representations by a simple translation procedure. In this respect we assume a similar approach to that which Copestake et al. (2005, page 312) and ERG take for tense. In ERG, tense is treated as a simple feature on events; there are no variables corresponding to, for example, time points. Copestake et al. argue that such a representation is preferable since it simplifies parsing and generation, and it suffices if it is possible to derive the deeper representation from this representation. Consider our analysis of the MP *into Paris* in Example (64). The meaning of $in_loc(l, paris)$ is that the localizer function corresponding to *in* maps *paris* to l . This localizer function is what we saw as IN of type $e \rightarrow (i \rightarrow (r \rightarrow t))$ in Kracht's formalism in Section 3.1. Hence Example (64) corresponds directly to Example (65) in Kracht's formalism. This is equivalent to Example (66), which is the analysis according to Kracht. If we then use Kracht's definition of IN in terms of

the local relation i this can be spelled out as Example (67), and if we use the definition for CF this can be translated further into Example (68).

$$(64) \ h: \text{cf_mod}(e, l), \ h: \text{in_loc}(l, \text{paris})$$

$$(65) \ \text{CF}(L)(e) \wedge \text{IN}(\text{paris}) = L$$

$$(66) \ \text{CF}(\text{IN}(\text{paris}))(e)$$

$$(67) \ \text{CF}(\lambda t \lambda r [i(\text{LOC}(\text{paris})(t), r)])(e)$$

$$(68) \ \exists t(\text{TIME}(e)(t) \wedge \neg i(\text{LOC}(\text{paris})(t), \text{LOC}(\mu(e))(t))) \wedge \\ \forall t(\text{TIME}(e)(t) \rightarrow \exists t' (\text{TIME}(e)(t') \wedge t \leq t' \\ \wedge i(\text{LOC}(\text{paris})(t'), \text{LOC}(\mu(e))(t'))))$$

A similar translation can be applied in the case of the other mode relations, including static. This means that from our simpler surface representations, we can construct Kracht's representations for application purposes when it is desirable.

The location variable l is unbound in the MRS representations. As with event variables in MRS, we assume an implicit quantifier for each location variable. But in contrast to the event quantifier, we assume that this quantifier has a fixed narrow scope, that is, it scopes over the modalizer and localizer relations, and is outscoped by all quantifiers binding referential indices, including the quantifier introduced by the complement of the PP. Moreover, this quantifier will be a quantifier claiming that there exists exactly one such region. Though the quantifiers are not explicitly represented, the translation into Kracht's functional language proceeds as if they were there.

One could ask why we bother to decompose the preposition at all. Why do we choose to analyze Example (61) as Example (62) when we could use the simpler Example (69).

$$(69) \ \langle h0, h1, \{h1: \text{under}(e, x)\}, \{\} \rangle$$

There are several reasons why we want to decompose. First, a preposition like *under* can be interpreted in any of three ways $\{h1: \text{st_mod}(e, l), h1: \text{under_loc}(l, x)\}$, $\{h1: \text{cf_mod}(e, l), h1: \text{under_loc}(l, x)\}$, or $\{h1: \text{tr_mod}(e, l), h1: \text{under_loc}(l, x)\}$. A simple translation procedure would have to propose all these three when the preposition is translated into a deeper representation. But by having the mode explicitly represented, we can to a certain degree disambiguate between the different modes during parsing. As this disambiguation is performed during the syntactic analysis, we believe that the disambiguated structure should also be reflected in the semantic representation. Second, we will see in Section 5.2 that the decomposition simplifies the interpretation of sentences where several prepositions co-occur. Third, we find anaphoric uses of locative expressions, and an explicit representation is needed to describe the sentence in Example (70).

$$(70) \ \text{Kim traveled to Paris, and Sandy lives there.}$$

Here, it is clear that the anaphoric *there* cannot point to the set of events denoted by the MP *to Paris*. And as we want *there* to denote a location rather than an entity, we cannot simply let it point to *Paris* either, unless we coerce *Paris* to a location. Rather, we let *there* point to the location corresponding to the LP level of the PP *to Paris*.

5. Norwegian Locatives

In this section, we will use Kracht's (2002) theory for two purposes. We will first show how locative prepositions can be classified along the modalizer and localizer dimensions, and secondly use the theory to explain differences between Norwegian and English locative prepositions. In Norwegian, we find no direct counterparts to the English cofinal prepositions *into* and *onto*. Instead, one uses a combination of an intransitive locative preposition, for example, *inne* (*inside_{st}*) or *inn* (*inside_{cf}*), followed by the relevant preposition, for example, *i* (*in*) or *på* (*on*). This was exemplified in the sentences in Examples (11) and (12). We return to our analysis of these constructions in Section 5.2. Thus, this section will focus mostly on the analysis of Norwegian locatives.

5.1 Classes of Locatives

Looking at the selectional patterns of Norwegian prepositions, we identify three distinct patterns:

1. **PP transitive.** These take PP complements. The most frequent are *fra* (*from*) and *til* (*to*), with *via/om* (*via*) and *mot* (*toward*) as possible candidates. Note that all these prepositions also occur with NP complements.
2. **NP transitive.** These take NP complements. All ordinary prepositions fall into this class.
3. **Intransitive.** These take no complements, and are semantically equivalent to a full PP with a contextually determined location. Examples are *her* (*here_{st}*), *hit* (*here_{cf}*), *inne* (*inside_{st}*), *inn* (*inside_{cf}*), and *innenfra* (*from inside*).

Furthermore, the analysis we propose is decompositional, and we classify locatives semantically along the modalizer and localizer dimensions. The modalizer dimension seems to have an upper boundary, as Kracht (2002) claims to have found evidence for five modalizers. The localizer dimension has in principle no upper boundary. In Table 1, we show how the different types of locatives are distributed over the localizer and modalizer dimensions. Prepositions in the same column share modalizer, and prepositions in the same row share localizer. For each Norwegian locative, the corresponding English locative is listed in parentheses.

Recall that prepositions can function as heads of MPs and LPs. For Norwegian and English, there is a **general correspondence** between the prepositions being heads of **static MPs** and being **LPs**. It is also worth noting the **transparent morphology** the intransitive prepositions show with respect to mode. The system is made up of a stem and a suffix, as follows: *-e* (static), *-efra* (coinitial), *-* (cofinal), *-om* (transitory) and *-over* (approximative).

5.2 More on the Internal Structure of Locatives

We follow Kracht (2002) in assuming that locatives have an MP structure of the form [M [L NP]]. Syntactically, this can be realized in several different ways. Relating the classes of prepositions to the MP structure above is done as shown subsequently. The underlined constituents of the MP structure correspond to the preposition of the class.

Table 1
Locatives classified with respect to mode.

	st	ci	cf	tr	ap
<i>PP transitive</i>	–	fra (from)	til (to)	via (via)	mot (toward)
<i>NP transitive</i>	ved (at) på (on) i (in) under (under)	fra (from) av (off) ut av (out of) –	til (to) på (onto) i (into) under (under)	via (via) over (across) gjennom (through) under (under)	mot (toward) – – –
<i>Intransitive/ Alternating^a</i>	her (here) inne (inside) hjemme (at home)	herfra (from here) innenfra (from inside) hjemmefra (from home)	hit (here) inn (inside) hjem (home)	– innom (–) hjemom (–)	hitover (toward here) innover (inward) hjemover (homeward)

^a See Section 5.2 for alternating prepositions.

The relevant part of the semantics for the locative expression is also shown, using the example sentences from Section 1.

1. **PP transitive.** These are analyzed as modalizers. They select for LP complements, which in turn contribute with the localizer relation in order to get at the full semantics of the locative.

$\underline{M} + L + NP$

(7) [A mouse appeared *from* under the table]' =
... ci_mod(*e, l*) ...

2. **NP transitive.** These are analyzed as combined modalizers and localizers. The lexical preposition is decomposed to get a uniform analysis of the locative PPs, and to facilitate a disambiguation of the modalizer part as explained at the end of Section 4.2.

$\underline{M} + \underline{L} + NP$

(3) [The mouse ran *under* the table]' =
... st_or_cf_or_tr_mod(*e, l*) \wedge under_loc(*l, x*) ...

3. **Intransitive.** These are analyzed as a modalizer and a localizer. The localizer function is special in the sense that the landmark is contextually determined. We have encoded this as a localizer relation taking only one semantic argument, of type location.

$\underline{M} + \underline{LP}$

(15) [Musa kom *derfra*]' =
... ci_mod(*e, l*) \wedge der_loc(*l*) ...

As mentioned, some intransitive prepositions may occur with a complement, as we saw in Examples (11)–(14). We name these **alternating locatives**, and analyze them as prepositions with an optional complement. When the alternating prepositions occur with a complement, the mode relation scopes over both the contextually determined

location and the LP complement location. This analysis is based on the fact that the alternating preposition introduces the mode, and does not merely disambiguate between modes, as the cofinal reading of *over* is only available when preceded by an alternating preposition, as shown in Examples (13) and (14). Thus, we end up with the following analysis of alternating prepositions:

4. **Alternating.** These are analyzed as a modalizer and a localizer, but take a second optional LP complement, otherwise identical to intransitive prepositions. Examples are *inn* (*inside_{cf}*), *inne* (*inside_{st}*), and *ute* (*outside*).

$\mathbf{M} + \mathbf{LP} + \mathbf{LP}_{\text{optional}}$

(12) [Musa løp *inn* i hullet]' =
...*cf*_mod(*e*, *l*) \wedge *inne*_loc(*l*) ...

The MRS for the LP complement analysis of Example (12) is given in Example (71) (where we only show the elementary predications and omit the quantifiers). Note that the sentence in Example (12) is syntactically and semantically ambiguous in our grammar. It also has a reading where the PP *i hullet* is not a complement to *inn*, but where both PPs modify the main event. The two different types of readings are shown in Section 6.6.

(71) ... *h4*: mus(*x*), *h5*: løpe(*e*, *x*), *h5*: *cf*_mod(*e*, *l*),
h5: *inne*_loc(*l*), *h5*: *i*_loc(*l*, *y*), *h8*: hull(*y*) ...

6. Implementation

In this section, we first describe the framework the grammar was implemented in. We then proceed to describing the classes of locative prepositions described in Section 5. The implementation of Kracht's (2002) emptiness principle will be central to our grammar.

6.1 Framework for the Implementation

MRS is a general format for representing different formal object languages and may be combined with different syntactic models. It has in particular been combined with HPSG grammars where it has been used in the large English Resource Grammar (Flickinger 2002). HPSG (Pollard and Sag 1994) is a unification-based non-transformational theory, using multiple inheritance type hierarchies and unification of typed feature structures as central formal mechanisms. The analysis we have presented so far is compatible with different syntactic models. We have made a small fragment of an HPSG grammar to illustrate the analysis. This fragment was implemented in the Linguistic Knowledge Builder (LKB) system (Copestake 2002). The LKB system is an implementation environment for writing HPSG grammars. The grammar was based on the Matrix "starter kit" for building HPSG grammars (Bender, Flickinger, and Oepen 2002). The Matrix contains types for building MRSs as the semantic content of a sentence compositionally. Principles restricting the syntax-semantics interface, as described by Copestake, Lascarides, and Flickinger (2001), are also encoded in the Matrix. For instance, a sign (phrase or word) only exposes its handle, main internal argument, and external argument to the rest of the grammar, such that, for example, a verb cannot control the arguments of its complement. The description of the grammar

presented here deviates somewhat from the Matrix, in order to improve readability. We also only present the relevant features of the grammar, as the actual descriptions of the signs in the grammar are quite complex, having to deal with many other aspects of the grammar. We refer to Jørgensen (2004) for details of the implementation.

6.2 Goals

We want our implementation to give the correct syntactic and semantic (MRS) analyses, while having minimal redundancy in the lexicon. We furthermore want to under-specify ambiguities wherever possible. Corresponding loosely to questions Q1–Q3 in Section 2, we need to consider the following dimensions that serve as the guidelines for the grammar presented here:

1. **Modalizer+Localizer/Localizer:** Locative prepositions must have the ability to denote sets of events (i.e., serving as M+L) and denote locations (L), depending on the syntactic context.
2. **Adjunct/Complement:** Locative prepositions must have the ability to serve as modificational adjuncts and as referential complements, depending on the syntactic context.
3. **Mode ambiguity:** Mode ambiguous locative prepositions should display this ambiguity in the semantics, and only be disambiguated semantically when the syntactic context allows it.

6.3 NP Transitive Prepositions

For sentence (3), we want to obtain the MRS in Example (72). The HPSG representation for the sign *under* is given in Figure 1. The bag of EPs provided by *under* is represented as a list which is the value of the MRS.RELS, and where, for example, the EP ‘*h5: under_loc(l, y)*’ is represented by an AVM where *under_loc* is the value of PRED, *h5* the value of LABEL, *l* is the value of ARG0, and *y* the value of ARG1. Moreover, *h5* is the local top of this MRS and that is encoded as the value of MRS.HOOK.LTOP.

(72) [The mouse ran under the table]’ =
 ⟨*h0, h5, {h1: def(*x, h2, h3*),
 h4: mouse(*x*), h5: run(*e, x*),
 h5: st_or_cf_or_tr_mod(*e, l*), h5: under_loc(*l, y*),
 h9: def(*y, h6, h7*), h8: table(*y*)}*⟩,
 {*h2 =_q h4, h6 =_q h8*}⟩

We arrive at the MRS in Example (72) as follows:

- **Mode underspecification:** The lexical entry for *under* introduces a disjunctive modalizer relation (*st_or_cf_or_tr_mod_rel*).
- **Combined M+L:** *Under* introduces two semantic relations (in MRS.RELS), one *mod_rel* (*st_or_cf_or_tr_mod_rel*) and one *loc_rel* (*under_loc_rel*).
- **Event modification:** The MOD contains a list of elements *under* may modify. We see from Figure 1 that *under* modifies a verb, and that the main variable of the verb (MOD<MRS.HOOK.INDEX>) is coindexed with

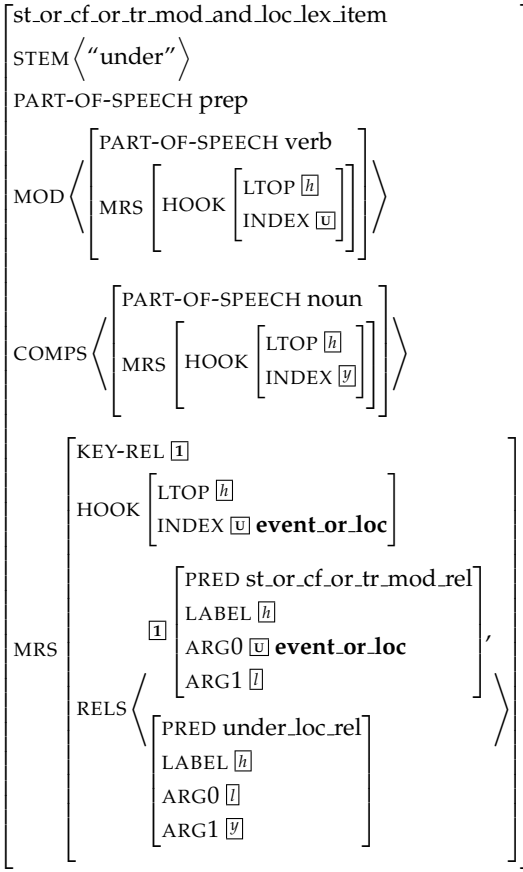


Figure 1 Simplified AVM for the NP transitive preposition *under*.

the main variable of the preposition (MRS.HOOK.INDEX), treating the preposition as an intersective modifier. (Disregard the underspecified variable type *event_or_loc* for now, except for the fact that because the verb’s MRS.HOOK.INDEX is an **event**, the underspecified type will also be constrained to be an **event** in cases of adverbial modification.)

6.4 PP Transitive Prepositions

We now turn to the sentence in Example (7), for which we want to achieve the MRS in Example (73).

- (73) [A mouse appeared from under the table]’ =
 - $\langle h0, h5, \{h1: \text{some}(x, h2, h3),$
 - $h4: \text{mouse}(x), h5: \text{appear}(e, x),$
 - $h5 : \text{ci_mod}(e, l), h5 : \text{under_loc}(l, y),$
 - $h9: \text{def}(y, h6, h7), h8 : \text{table}(y)\},$
 - $\{h2 =_q h4, h6 =_q h8\}\rangle$

This is more complex, as we want to arrive at this MRS while still using the sign for *under* from Figure 1. The representation for *from* acting as a PP transitive preposition is given in Figure 2. We have two different entries for the preposition *from* in the lexicon, one where it is PP transitive and one where it is NP transitive. The entry for the NP transitive *from* is more or less identical to *under*, except that it introduces different semantic relations (*ci_mod_rel* and *at_loc_rel*).

The two signs in Figure 1 and Figure 2 together correspond to our implementation of the emptiness principle, by unification of the constraint *id_rel* on COMPS<MRS.KEY-REL> in Figure 2 and MRS.KEY-REL in Figure 1. As we recall from Section 2.7, the emptiness principle claims that when an element is fixed regardless of the interpretation of the sentence, it is interpreted as the identity relation. Kracht’s claim is more radical than our implementation, as he claims that *any* element may be empty. We have only implemented the emptiness principle for prepositions.

The MRS in Example (73) is obtained as follows:

- **Identity relations:** The claim that any preposition can be semantically empty corresponds in our type hierarchy to claiming that every relation introduced by a preposition has an identity relation as a subtype of it, as shown in Figure 3. Recall that empty elements must be syntactically selected. Thus we constrain the selecting head, whether a verb or a modalizer, to select a preposition with an empty element.
- **Localizer–Static modalizer correspondence:** As already noted, there is a general correspondence between the static prepositions and the prepositions that figure as complements of PP transitive prepositions.
- **Selectional constraints:** *From* selects a static preposition whose key semantic relation is the identity relation. This is represented in MRS.KEY-REL, a feature that every sign exposes according to the

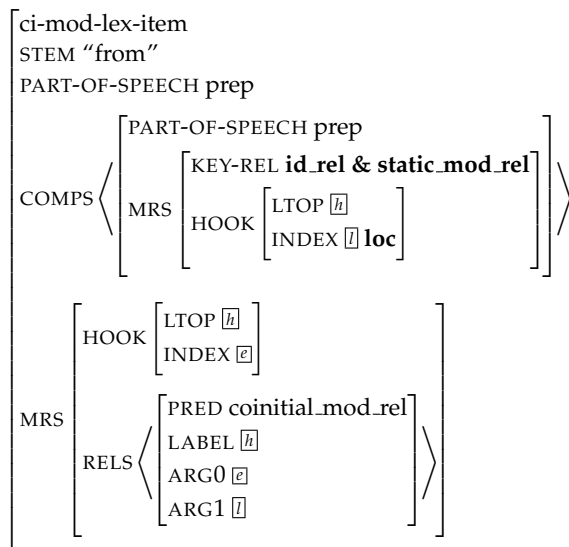


Figure 2
The PP transitive preposition *from* and LP selection.

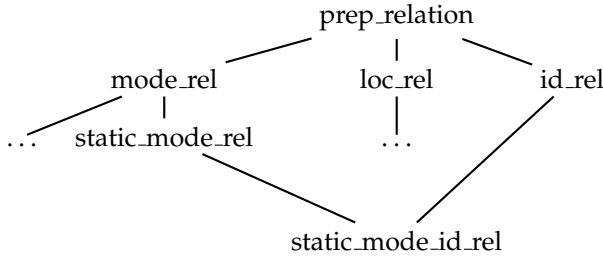


Figure 3 A type hierarchy of prepositional relations.

principles for the Matrix (Bender, Flickinger, and Oepen 2002). MRS.KEY-REL is constrained to be of a common subtype of static_mod_rel and id_rel, represented in the AVM by '&'.

- **Type constraint:** The MRS.HOOK.INDEX, i.e., the external argument of the complement, is constrained to be of type **loc**. Recall from Figure 1 that this was underspecified as **event_or_loc**.

6.5 Intransitive Prepositions

Intransitive prepositions are implemented similarly to NP transitive prepositions, except that they take no complements (empty COMPS list), and that they introduce a loc_rel with only one variable. The MRS we obtain is shown in Example (74).

(74) [A mouse ran here]' =
 ⟨h0, h5, {h1: some(x, h2, h3),
 h4: mouse(x), h5: run(e, x),
 h5 :st_or_cf_mod(e, l), h5 :here_loc(l)},
 {h2 =_q h4}⟩

6.6 Alternating Prepositions

Alternating prepositions are implemented as something between a PP transitive preposition and an intransitive preposition. They take optional static PP complements, which in turn are interpreted as LPs (identical to LP selection for PP transitive prepositions). The location variable of the internal loc_rel and the complement loc_rel are coindexed, such that the mod_rel predicates over both localizer relations. The MRSs for the syntactic analysis of Example (10), where the second PP is a complement of down, is shown in Example (75). As mentioned in Example (12) in Section 5.2, Example (10) is syntactically and semantically ambiguous in our grammar. It has a reading where the PP under the bridge is not a complement to down, but where both PPs modify the main event. This is shown in Example (76).

(75) [A child ran (MP down (LP under the bridge))]' =
 ⟨h0, h5, {h1: some(x, h2, h3), h4: child(x), h5: run(e, x),
 h5 :cf_mod(e, l), h5 :down_loc(l), h5 :under_loc(l, y),

Downloaded from http://direct.mit.edu/col/article-pdf/35/2/229/1798610/col.06-69-prep5.pdf by guest on 21 May 2022

$h6: \text{def}(y, h7, h8), h9: \text{bridge}(y)\},$
 $\{h2 =_q h4, h7 =_q h9\}$

- (76) [A child ran (_{MP} down) (_{MP} under the bridge)]' =
 $\{h0, h5, \{h1: \text{some}(x, h2, h3), h4: \text{child}(x), h5: \text{run}(e, x),$
 $h5 : \text{cf_mod}(e, l1), h5 : \text{down_loc}(l1), h5 : \text{st_or_cf_mod}(e, l2), h5 : \text{under_loc}(l2, y),$
 $h6: \text{def}(y, h7, h8), h9: \text{bridge}(y)\},$
 $\{h2 =_q h4, h7 =_q h9\}\}$

6.7 PP Transitive Verbs

The mechanisms we introduce for LP selection and mode underspecification have until now been used to describe complex locative PPs. But the same mechanisms can be used for other heads selecting PP complements, for example, verbs. Recall Bierwisch's (1988) claim that *liegen* (*lie*) is a relation between the subject and a location. This can now be implemented directly, by stating in the description of the verb that it selects an LP complement, identical to the complement of the PP transitive (modalizer) prepositions.

More interesting is the class of verbs known as put-verbs. Such verbs take both static and cofinal complements, and express directionality in both cases as we saw in Example (8). Kracht's two-layering enables us to let the verb select the full PP (MP) as the syntactic argument, but only the denotation of the LP as its semantic argument. The verb can thus express the movement, while the PP contributes the goal of the movement to the semantics. We encode this by describing the PP complement of the verb to be an LP (denoting a location) of static or cofinal mode (in $\text{COMPS} \langle \dots, [\text{MRS.KEY-REL}] \rangle$). Thus, we arrive at the same semantics for both sentences in Example (8). At the same time we constrain the verb to only select PPs of static or cofinal mode. A simplified AVM for the verb *put* is shown in Figure 4. A similar treatment can be given to any verb that semantically selects a location rather than an event as the denotation of the PP complement.

6.8 Summary

We have achieved our goals for the implementation from Section 6.2 as follows:

1. **Adjunct/Complement:** All locatives have the ability to serve as modificational adjuncts through the MOD list. All locatives have the ability to serve as complements, by being selected by a head, for example, a modalizer (*from*) or a verb.
2. **Modalizer+Localizer/Localizer:** When serving as a modificational adjunct, the item on the MOD list constrains the MRS.HOOK.INDEX of the PP to be an event, namely, the PP denotes a set of events. When serving as a complement, the selecting head may constrain the MRS.HOOK.INDEX of the PP to be a location, and at the same time specify the main semantic relation, the MRS.KEY-REL, to be an *id_rel*, which is a subtype of all *mod_rels*.
3. **Mode ambiguity:** We introduce disjunctive semantic relations for *mod_rels*. These may be further constrained by the syntax, for

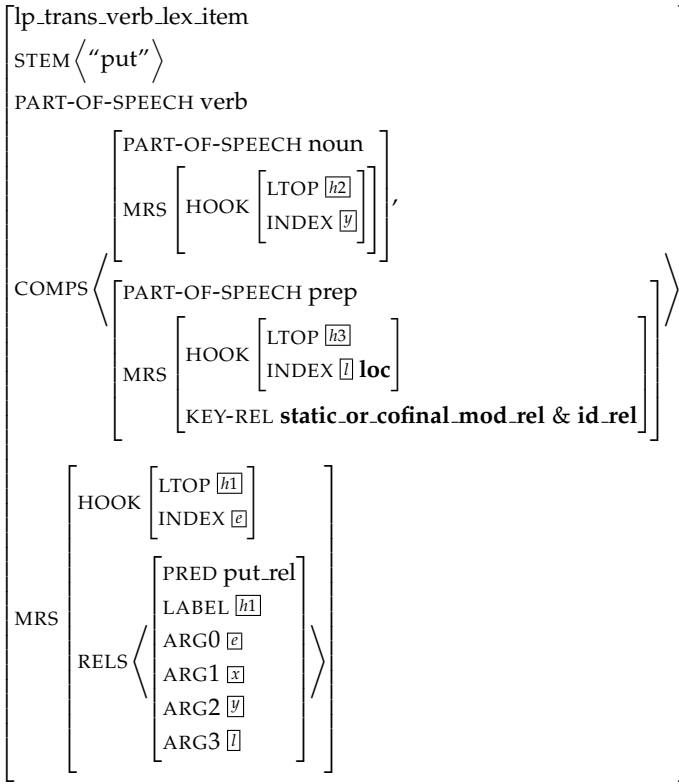


Figure 4
Simplified AVM for the verb *put*.

example, by specifying that only motion verbs may be modified by directional modes. This particular feature has not been implemented yet, however.

7. Machine Translation

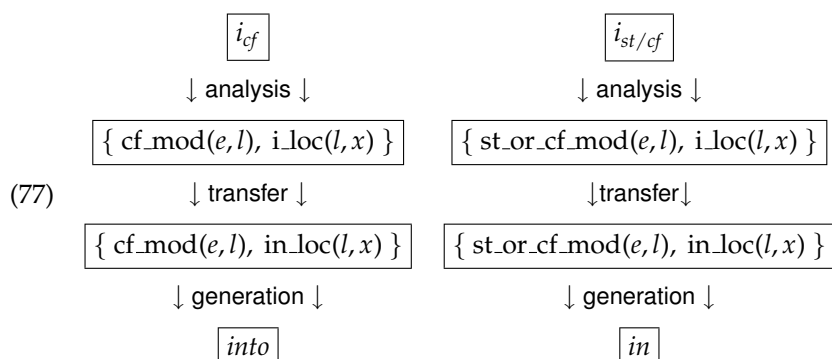
Our analysis can be applied in several types of applications. First of all, it serves as the basis for a coherent analysis of locatives, for example, in that PP complements of both prepositions and verbs are analyzed in a uniform manner. It also makes the syntax–semantics interface clean, in implementing the dualistic nature of locatives as referential and modificational. We believe one of the strengths of this analysis concerns the choice of the level of representation. In an NLP system where spatial reasoning is required, one may choose to spell out the details in the analysis, as shown in Section 4.2. In other NLP applications, the superficial representation is sufficient, and one does not want to introduce a complex model involving movers, time points, and parameterized neighborhoods. In this section, we show how our analysis may serve as the best level for representations in a transfer-based machine translation system, due to the interlingual nature of the modes.

Kracht (2002) argues that the modes can be found across a wide range of languages, making them good candidates for true interlingua predicates. The localizer

functions, however, are language-dependent, and there exist different systems for locatives in different languages. Also, whether the particular combination of modalizer and localizer is lexicalized is language-specific, as we see in Section 7.2. We will illustrate the flow in a semantic-based transfer MT system from Norwegian to English, with two examples. The modes are treated as interlingua predicates, whereas the localizer functions are language-dependent, and thus transferred by the transfer module. In Section 7.1, we consider an NP transitive preposition, and show how a mode-disambiguated preposition may translate differently from the mode-ambiguous preposition. And in Section 7.2, we describe how an intransitive preposition is translated into a complex PP consisting of a PP transitive preposition and an intransitive preposition.

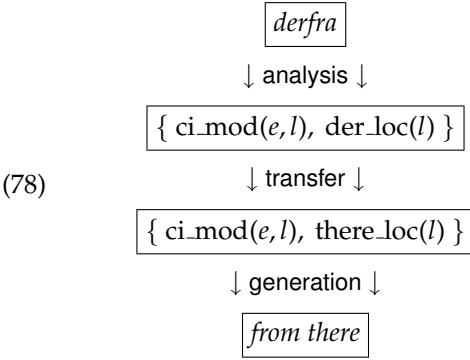
7.1 Translation of an NP Transitive Preposition

Consider the Norwegian preposition *i* (*in*). It is ambiguous between a static and a cofinal reading, but may be disambiguated by the syntactic context. The translation flow of the cofinal and ambiguous readings of the locative *i* (*in*) are shown in Example (77).



7.2 Translation of an Intransitive Preposition

Intransitive locatives, or context-dependent locatives, provide more interesting mismatches between Norwegian and English locative expressions. In Table 1 from Section 5.1, we see that Norwegian has a rich inventory for expressing context-dependent spatial relations. One example is the Norwegian locative *derfra*, as seen in Example (15), which translates to the English complex *from there*. These kinds of mismatches are now easily explained. In Norwegian, we find a lexicalization for the combination of the cointial mode and the localizer function *der_loc* (semantically interpreted as distal, and contrasted by proximal/close). In English, we do not find this lexicalized preposition, and the two semantic relations (modalizer and localizer, respectively) are realized as separate lexical items. The translation flow of the locative *derfra* is shown in Example (78).



Translating in the other direction, we also receive a second possible translation: *fra der*. This can be blocked, however, if the transfer mechanism allows for rule ordering where rules can be optional or obligatory, as in the LOGON system (Oepen et al. 2004). Then the rule choosing *derfra* is tried first and if it succeeds the other option is blocked.

8. Other Computational Approaches

In Section 2 we considered different linguistic approaches to locatives that were not computational. We will here compare our approach to earlier computational approaches with similar goals to ours.

8.1 Creary et al.

Creary, Gawron, and Nerbonne (1989) describe a fully implemented system for database queries. They propose to take locative PPs to be arguments rather than modifiers. They observe that such an approach must explain occurrences of several PPs as in Example (79).

(79) Al works on Mass Ave in Boston near MIT.

Their solution is to let each PP denote a region. The denotations of the three PPs are intersected (as point sets) resulting in one region satisfying all three of them. This region is then taken as an argument to the semantics of the verb.

Such a solution seems to work for this example. But all the examples in the article are static PPs. This approach does not so easily extend to directional prepositions because directional prepositions cannot be taken to denote regions.

8.2 The Core Language Engine

In the Core Language Engine, ordinary PPs are considered modifiers (Alshawi 1992). Semantically, an ordinary preposition is a relation between two arguments, the landmark object, and the event denoted by the verb in the case of verbal modification. There is no special treatment of locative prepositions and no distinction is made between directional and static prepositions. This is in many ways a baseline approach to prepositions in event-based computational grammars, and we sketched it in Example (69). In addition to the ordinary PPs, four other classes of PPs are recognized where one is the class of subcategorized prepositions which are semantically empty, as in *handed a book to Mary*.

Downloaded from http://direct.mit.edu/col/article-pdf/35/2/229/1798610/col.06-69-prep5.pdf by guest on 21 May 2022

We argued in Section 4.2 for why we think our refined analysis is an improvement over the baseline approach.

8.3 VerbMobil

In VerbMobil the semantics of prepositions are also considered relations between the landmark and the event (or object) it modifies (Buschbeck-Wolf 1995). Syntactically, directional PPs may be complements, but they are still considered semantic modifiers. In addition to relations and arguments, the semantic representations in VerbMobil are decorated with additional information which is exploited in the transfer step of the translation. In particular, the representations for PPs and the lexical entries for prepositions contain a feature *ROLE* as part of the argument (to be) filled by the landmark. This *ROLE* can contain information like *way_rel* for directionality, and its negation for static PPs. In German, this property will correspond to whether the case of the landmark is accusative or dative.

Though this extra *ROLE* information may be a useful tool in translation, it lacks a denotational meaning. Moreover, this approach does not explain how we can make reference to locations, nor the structure of the PP transitive prepositions.

8.4 Trujillo's Spatial Prepositions and Machine Translation

Trujillo (1995) analyzes locatives in the context of transfer-based machine translation. He first develops a general format for a flat semantic representation called **indexed lexeme lists**, and shows how transfer between English and Spanish can be carried out on such representations. He then investigates differences in properties between English and Spanish spatial prepositions, and discusses the proper semantic representation of locatives and how this can be exploited in translation.

The semantic representations used by Trujillo are quite similar to MRSs and introduce many properties that can also be found in the MRS formalism. One difference between the two is that indexed lexeme lists lack an equivalent to the scoping mechanism in MRS, and Copestake et al. (2005) show how this can be a problem for semantic disambiguation and translation. The MRS formalism can be considered an improvement over the indexed lexeme lists.

Trujillo does not decompose the prepositions. Syntactically, both static and directional prepositions are considered adjuncts and semantically they are considered modifiers. Prepositions denote relations between the event denoted by the verb and the landmark object. Trujillo observes that prepositions can be stacked as in Example (7) from the introduction, repeated as Example (80). He proposes an analysis where the relevant parts are shown in Example (81) in a notation comparable to ours.

(80) The mouse appeared from under the table.

(81) mouse(x), table(y), appear(e, x), from(p, e, q), under(q, p, y)

As we see, Trujillo lets the prepositions take three arguments. To facilitate a comparison with our implementation and with the English Resource Grammar in the next section, we will call the three arguments ARG0, ARG1, and ARG2, respectively; for example, p is

the ARG0 of *from*(p, e, q), e is the ARG1, and q is the ARG2.⁷ The last two arguments, ARG1 and ARG2, are filled with the event being modified and the landmark, respectively, as in the Core Language Engine approach. When the two prepositions are combined, the lower preposition, *under*, modifies the object which is the ARG0 of *from*, and at the same time the object which is the ARG0 of the lower preposition is taken to be the landmark, the ARG2 of *from*.

Though this solution might work technically for translation purposes, we have problems making sense out of the status of these objects. Trujillo (1995, page 172) says that “these correspond to the place objects that Sondheimer (1978) proposes, although for the purposes of this thesis they become place holders for spatial relations.” The ERG also uses a representation with some of the same properties, and we will return to a more detailed comparison to our proposal in the following section. Trujillo does not decompose the prepositions, but he classifies the prepositions with respect to different semantic properties including whether they are static or directional and different directionality types, and places them in a semantic type hierarchy. We think that such a type hierarchy might be useful for selecting the translations of the locative part of a preposition, but from our perspective, this belongs to the lexical semantic properties and is beyond the scope of this article.

8.5 The English Resource Grammar

The English Resource Grammar (Flickinger 2002)⁸ is similar to our proposal in using MRS as the semantic formalism in a linguistic-based computational grammar, but the treatment of locatives is different from ours. Because the ERG treatment of prepositions is quite elaborate, treating different prepositions in different ways and interleaving the semantic analysis with the grammar, we will not be able to give a full and fair presentation of the full ERG analysis, but we will try to sketch some of its main properties regarding locatives, and compare them to our approach.

8.5.1 Prepositions in ERG. In Example (82) we see a simplified logical representation of the relevant aspects of the MRS representation ERG ascribes to Example (1). The preposition introduces a relationship between its ARG1, the main event of the sentence, e_1 , and its ARG2, the landmark *Paris*. In addition, prepositions, like verbs, introduce their own events as an ARG0. This event is used as the main event of sentences where the PP is used predicatively (*Kim is in Paris*), but e_2 is not used in Example (82). An intransitive preposition like *down* also modifies the main event and introduces its own event, as seen in the ERG analysis of Example (9), shown in Example (83). This analysis is similar to the ERG analysis of intersective adverbs, like *fast*. Syntactically, a locative PP can be adjoined to any kind of VP and modify the VP event. In addition, there is a class of directional verbs, like *go*, *run*, and *drive*, which can take a locative PP as an argument. Semantically, however, this is interpreted as modification. Thus even though sentence (2) gets two different syntactic analyses, they are both ascribed the same semantics (see Example (84)).

(82) $\text{sleep}(e_1, \text{kim}) \wedge \text{in}(e_2, e_1, \text{paris})$

⁷ Trujillo used other names on the features in his implementation.

⁸ See <http://www.delph-in.net/erg>. ERG has developed during the time we have worked on this article. We will here compare to the version of 17-mar-07.

- (83) $\text{some}(x, \text{child}(x), \text{run}(e_1, x) \wedge \text{down}(e_2, e_1))$
- (84) $\text{drive}(e_1, \text{kim}) \wedge \text{into}(e_2, e_1, \text{paris})$
- (85) $\text{mouse}(x) \wedge \text{run}(e_1, x) \wedge \text{under}(e_2, e_1, y) \wedge \text{table}(y)$
- (86) $\text{mouse}(x) \wedge \text{run}(e_1, x) \wedge \text{under}_{\text{dir}}(e_2, e_1, y) \wedge \text{table}(y)$

Only directional PPs can be arguments to directional verbs. There are three different types of locative prepositions in ERG. **Non-directional** prepositions, like *at*, can head a PP which is adjoined to a VP, but cannot head a PP which is argument to a directional verb. **Ambiguous** prepositions, like *in* and *under*, can head PPs that are adjoined to verbs, as well as PPs that are arguments to directional verbs. The two different syntactic analyses the grammar ascribes to Example (3) are semantically both considered as event modification, but the semantic predicate ascribed to the preposition is different in the two cases. When the PP is argument to the verb the preposition is constrained to be directional (Example (86)), but not when it is adjoined to the VP (Example (85)). The third class of prepositions, represented by *to*, *into*, and *onto*, which we would classify as **directional**, can head PPs in both types of constructions, but ascribe identical semantics to the two (Example (84)).

The main differences between ERG and our proposal so far are the following. ERG does not decompose prepositions semantically. ERG introduces a distinction between directional and static, but does not distinguish between different modes of directionality. In ERG there is a correspondence between syntax and directionality, as PPs that are syntactic arguments of motion verbs have to be read directionally. On the other hand, unambiguously directional prepositions are also allowed to be non-arguments and modify any VP. In our analysis, a semantically modifying PP is syntactically always adjoined to the verb. We do not postulate a syntactic ambiguity in Example (2). And even in semantically ambiguous sentences, like Example (3), we assign only one syntactic structure and keep the mode underspecified in the syntax–semantics interface.

8.5.2 Put-Verbs. Example (8) gets the analysis in ERG shown in Example (87). The main point of the analysis is that *put* takes two individual arguments, corresponding to the subject NP and the direct object NP, and one proposition, which again describes a spatial relationship between two individuals corresponding to the direct object NP and the complement of the locative preposition.

- (87) $\text{def}(y, \text{book}(y), \text{def}(z, \text{table}(z), \text{put}(e_1, \text{kim}, y, \text{on}/\text{onto}(e_2, y, z))))^9$
- (88) $\text{def}(y, \text{book}(y), \text{def}(z, \text{table}(z), \text{put}(e_1, \text{kim}, y, l), \text{on}(l, z)))$

The most likely way to read this semantic representation when the PP is *on the table* must be as a resultative. The result of the action is that the relationship *on* holds between the book and the table. Exactly the same type of analysis is applied in ERG when the PP is

⁹ We have simplified the representation a little, by not including the so-called propositional messages (*prpstn.m*), which is part of the ERG version we compare to, as they are without importance for the discussion.

the unambiguously directional *onto the table*, that is, now *put* is a relationship between the subject, the object, and the proposition expressing that the relationship *onto* holds between the book and the table. But what should it mean that the relation *onto* holds between the two objects? In our analysis, as discussed in Section 6.7 and illustrated in Example (88), the verb expresses a relationship between an object and a location directly. The sentence with *onto* ends up with exactly the same semantic representation as the sentence with *on*.

8.5.3 Alternating Prepositions. When an intransitive preposition is followed by a PP, as in Example (10), repeated here as Example (89), ERG provides two different analyses. The first is that the two locatives both modify the main event, e_1 , of the sentence independently of each other as in Example (90). This analysis corresponds to our analysis where two PPs modify the same VP (note Example (76) in Section 6.6), presented in logical form in Example (92). We read the second ERG analysis (Example (91)) as *down* modifying the running event. At the same time *down* introduces its own event e_2 , and this event is modified by *under the bridge*, in the same way as *under the bridge* would modify a verb.

(89) A child ran down under the bridge.

(90) $\text{some}(x, \text{child}(x), \text{def}(y, \text{bridge}(y), \text{run}(e_1, x) \wedge \text{down}(e_2, e_1) \wedge \text{under}(e_3, e_1, y)))$

(91) $\text{some}(x, \text{child}(x), \text{def}(y, \text{bridge}(y), \text{run}(e_1, x) \wedge \text{down}(e_2, e_1) \wedge \text{under}(e_3, e_2, y)))$

(92) $\text{some}(x, \text{child}(x), \text{def}(y, \text{bridge}(y), \text{run}(e_1, x) \wedge \text{cf_mod}(e_1, l_1) \wedge \text{down_loc}(l_1) \wedge \text{st_mod}(e_1, l_2) \wedge \text{under_loc}(l_2, y)))$

(93) $\text{some}(x, \text{child}(x), \text{def}(y, \text{bridge}(y), \text{run}(e_1, x) \wedge \text{cf_mod}(e_1, l_1) \wedge \text{down_loc}(l_1) \wedge \text{under_loc}(l_1, y)))$

We understand this as ERG recognizing the same ambiguity in Example (89) as we do, and that ERG tries to capture the reading we analyzed as Example (75), presented in logical form in Example (93), in this way. But it is neither obvious what kind of event is introduced by *down*, nor how this event differs from the event introduced by *run*. We think there is only one event involved here. In this event the child moves downward and ends up under the bridge. Hence it is a cofinal movement where the location is constrained both by *down* and *under the bridge*. In the first reading, where the two PPs modify the VP independently (Example (92)), there is still only one event, but because the two PPs might have different modes—e.g., *down* may be cofinal and *under the bridge* may be static—two different locations are considered. The cofinal *down* considers the location where the movement ends, while the static *under* considers the location of the whole event.

Another difference between our approach and ERG is that ERG ascribes the same two types of analysis to a sentence like Example (94).

(94) A child went down from the bridge.

We do not find this sentence ambiguous in the same way as Example (89). We think it can only mean that both PPs modify the VP. This is the only possible analysis in our approach, because *from the bridge* cannot be argument to *down*, as *down* in our grammar only allows static PPs as complements (interpreted as LPs), whereas *from* is unambiguously coinital.

8.5.4 PP Transitive Prepositions. Example (7) gets an analysis like Example (95) in ERG, and our analysis is presented in Example (96). The analysis in Example (95) resembles the one in Example (91), namely, *from* is treated as if it were intransitive and its last argument place is not filled; the variable u does not occur elsewhere. As with the analysis in Example (91), we have problems in understanding exactly how the event e_2 in Example (95) is different from the main event e_1 .

(95) $\text{some}(x, \text{mouse}(x), \text{def}(y, \text{table}(y), \text{appear}(e_1, x) \wedge \text{from}(e_2, e_1, u) \wedge \text{under}(e_3, e_2, y)))$

(96) $\text{some}(x, \text{mouse}(x), \text{def}(y, \text{table}(y), \text{appear}(e_1, x) \wedge \text{ci_mod}(e_1, l) \wedge \text{under_loc}(l, y)))$

8.5.5 Lexicalized Locatives. ERG also contains a quite different analysis for some locative and temporal prepositions and adverbs. We only consider the locative ones here. The adverb (intransitive preposition) *home* in Example (97) is analyzed in ERG as a modifier (adjective) on an individual y which is a definite *place* (Example (98)). In addition, it introduces the *unspec_loc*-relation which behaves as a preposition, taking the place y as its landmark (ARG2), modifying the main event of the sentence, e_1 (ARG1), and introducing its own unused event, e_2 (ARG0). Also the adjectival modifier *home* introduces its own unused ARG0 event.

(97) The child went home.

(98) $\text{def}(x, \text{child}(x), \text{implicit_q}(y, \text{place}(y) \wedge \text{home}(e_3, y), \text{go}(e_1, x) \wedge \text{unspec_loc}(e_2, e_1, y)))$

(99) $\text{def}(x, \text{child}(x), \text{go}(e_1, x) \wedge \text{cf_mod}(e_1, l) \wedge \text{home_loc}(l))$

ERG's lexicalized locatives (Example (98)) bear some similarities to our analysis (Example (99)) if one considers the individual place y in Example (98) to correspond to the location l . The *unspec_loc*-relation in ERG corresponds to the modalizer *cf_mod* in our analysis. But whereas the *unspec_loc*-relation is unrestricted and will represent any directionality type or static, the modes in our analysis can be constrained by the lexicon and syntax to only represent a set of possible modes. In this example the mode will be unambiguously co-final. A more principal difference between ERG and our analysis is that ERG introduces two quite different analyses of the locative intransitive prepositions (adverbs), the one exemplified by Example (98) and the analysis in Example (83) of Example (9). We ascribe the same type of analysis to Example (9) and Example (97).

8.5.6 *Consequences for Translation.* Our analysis, following Kracht, is meant to be universal across different languages. A locative expression always has the same form $st/ci/cf/tr/ad(e,l) \wedge \text{“localizer_relation”}(l,x)$. Which parts of the sentence contribute which parts of this representation might vary across languages. In a transfer-based MT system this means that the mode part will be interlingual, whereas the transfer rules will cater for the localizer. In spirit, this is quite similar to the treatment of temporal prepositions in the LOGON system (Oepen et al. 2004). On the ERG analysis, there is no reason to expect such a parallelism between two languages. What we have classified as an intransitive use of a locative preposition may be analyzed as either a simple intersective adverb or as a noun combined with the *unspec_loc*-relation in each of the two languages. This yields four different types of transfer rules, and which rule type to invoke for each word pair cannot be predicted without inspecting the relevant lexical entries in the grammars for the two languages.

Moreover, there is also a third type of analyses of intransitive uses of locatives in ERG, where the locative introduces an adjective, the noun *place* (*sted* in Norwegian), and an implicit quantifier just like the lexicalized locative in Example (98). But instead of the *unspec_loc*-rel, it introduces a full transitive preposition, like *from* or *to*. This is illustrated by Example (100) which in the Norwegian LOGON grammar is analyzed as Example (101).

(100) Barnet gikk herfra.
Child.DEF went here-from.
'The child went from here.'

(101) $\text{def}(x, \text{barn}(x), \text{def}(y, \text{sted}(y) \wedge \text{her}(e_3, y), \text{gå}(e_1, x) \wedge \text{fra}(e_2, e_1, y)))$

(102) $\text{def}(x, \text{child}(x), \text{implicit_q}(y, \text{place}(y) \wedge \text{here}(e_3, y), \text{go}(e_1, x) \wedge \text{from}(e_2, e_1, y)))$

In this case, such an analysis works well for transfer into English because it can be directly transferred into the representation for *from here* in Example (102). If we exchange *herfra* with *hit*, a similar analysis of Norwegian would exchange *fra* with *til* in Example (101). The most natural translation into English would be *The child went here*, and one would need yet another transfer type. Thus one might end up with nine different possible types of transfer relations between intransitive locative prepositions if both languages allow for the three different types of analyses.

8.6 NorGram

NorGram is a computational Lexical Functional Grammar for Norwegian. It has been extended to produce MRSs, as part of the LOGON system (Oepen et al. 2004).¹⁰ The Norwegian MRSs are used as input to transfer into English MRSs. NorGram follows the same design principles for MRS constructions for locatives as ERG, but departs from ERG in the analysis of PP transitive prepositions. If the NorGram analysis had been

¹⁰ See <http://www.emmtee.net> for the LOGON project.

applied to the English sentence in Example (7), we would get the analysis given in Example (103).

$$(103) \text{ some}(x, \text{mouse}(x), \text{ def}(y, \text{table}(y), \text{ appear}(e_1, x) \wedge \text{ from}(e_2, e_1, e_3) \wedge \text{ under}(e_3, u, y)))$$

On this analysis, the first preposition *from* is transitive, and it takes the PP *under the table*, or rather the event introduced by this preposition, as an argument. In contrast to ERG, this analysis allows the last argument of a preposition like *from* to be an event, e_3 , as well as an individual. Comparing this to our analysis in Example (96) and observing that e_3 plays a similar role as intermediary here as the location l does in our analysis, we can see some parallels, where $\text{from}(e_2, e_1, e_3)$ corresponds to $\text{ci_mod}(e_1, l)$, and $\text{under}(e_3, u, y)$ corresponds to $\text{under_loc}(l, y)$. Seen this way, the NorGram analysis of prepositions can be thought of as relating the landmark, its ARG2, to both an LP as its ARG0, and an MP as its ARG1. But the ARG0 and the ARG1 are never both used in the same analysis. This means that, technically, one could use the NorGram formalism as a surface representation for these types of constructions and translate into our representation at a deeper level. But such a translation would not be uniform; other correspondences would have to be applied for other constructions. Moreover, we think our representation is cleaner and conceptually clearer. All the introduced variables get used, and have a natural interpretation in our analysis. In particular, the variable connecting the two parts, l , is a location, whereas in the NorGram analysis, the connecting variable, is typed as an event, and it is not clear what kind of denotation could be given to this event.

9. Evaluation

Because we have so far only made a pilot implementation of the analysis, it is not possible to evaluate the grammar directly on a test corpus. The evaluation will therefore consist of two parts. First, a short corpus study of the LOGON Corpus, in order to examine the relevance of an analysis of locatives; and second, a description of the most important dimensions found in the data set, both as a summarization of our approach and as a brief comparison to ERG and Kracht, and also to construct a small test suite.

9.1 A Short Corpus Study: *The LOGON Corpus*

In order to evaluate the relevance of our analysis of locatives, we have investigated the distribution of prepositions in the LOGON Corpus. This corpus was developed as part of the LOGON project, where semantic MRS-based transfer is combined with stochastic ranking, translating texts from the hiking domain from Norwegian into English (Oepen et al. 2004).¹¹ The Norwegian part of the LOGON Corpus consists of 2,109 sentences and 30,348 words. There are 2–3 reference translations into English of the sentences, and ERG has been adopted to parse the English part. In the Norwegian part, 15.4% of the word occurrences (4,667) are prepositions. This number is slightly higher than in the Oslo Corpus¹² (13.6%), a balanced corpus of Norwegian texts.

11 See <http://www.emmtee.net> for the LOGON project, and also <http://www.hf.uio.no/tekstlab/prosjekter/tourist/index.htm> for the corpus.

12 <http://www.tekstlab.uio.no/norsk/bokmaal/>.

We have classified the prepositions as being locative or not, and find that about 50% of these (2,333) are locative. Furthermore, we have classified a subset of about one fifth of the locative prepositions with respect to syntactic role. We find that approximately 50% are adverbial, approximately 30% are adnominal, and the rest occur in other types of constructions, for example, predicative, or are difficult to classify. This preliminary corpus study indicates that our analysis of adverbial locatives are relevant for approximately 25% of the prepositions occurring in the LOGON Corpus.

9.2 Data Dimensions

As a basis for an evaluation of our analysis of locatives, we have gone through the data found in (Kracht 2002) and the LOGON Corpus. We present the most relevant dimensions herein, and compare them briefly to Kracht and ERG. In addition to motivating a short test suite for evaluation, this section is also intended to demonstrate the degree to which we cover the same fragment as Kracht does.

9.2.1 Adjunct Type. Locatives as adjuncts are either **adnominal** or **adverbial**. We analyze adverbial locatives according to the five modes. We do not cover adnominal locatives in our pilot implementation, and will therefore not include adnominal locatives in the test suite.

Kracht: Adverbial locatives are analyzed according to the five modes. Adnominal static locatives are treated similarly to static adverbials, but denote entities and not events.¹³

ERG: All locatives in the lexicon are underspecified with respect to denotation type, and are constrained by unification with the modified element to either an entity (adnominal) or an event (adverbial).

9.2.2 Directionality. **Directional locatives** can (in the general case) modify **directional or motion verbs** as in Example (t.2), but not **static verbs** as in Example (t.4). **Static locatives** can modify both classes of verbs, as in Examples (t.1) and (t.3). We specify on the directional prepositions that they only modify directional verbs, whereas static prepositions can modify all verbs.

- (t.1) Musa løp ved buret.
The mouse ran by the cage.
- (t.2) Musa løp til buret.
The mouse ran to the cage.
- (t.3) Musa satt ved buret.
The mouse sat by the cage.
- (t.4) *Musa satt til buret.
*The mouse sat to the cage.

Kracht: Four of the five modes are directional. All modes denote events, but the four directional modes predicate movement of an event participant, whereas the static

¹³ Directional adnominal locatives are also discussed (Kracht 2002, pages 199–202), but not treated formally.

mode locates the event itself. Directional verbs are defined indirectly, in that directional PPs only modify verbs denoting events that have an eligible mover.¹⁴

ERG: Directional prepositions are a subtype of ordinary prepositions. When selected by a directional verb, the preposition is constrained to be of the directional subtype. But all locatives, directional or not, can modify all VPs; hence, for example, Example (t.4) is accepted.

Note that the distinction between static and directional verbs is hard to make, and that some verbs, for example, manner of motion verbs, can be coerced to a directional reading when modified by an explicit directional locative as in Example (t.5). The directional reading is not available when modified by an ambiguous locative, however (Examples (t.6 and t.7)).

- (t.5) Musa danset inn på bordet.
The mouse danced onto the table.
- (t.6) Musa danset på_{st} bordet.
The mouse danced on_{st} the table.
- (t.7) *Musa danset på_{cf} bordet.
*The mouse danced on_{cf} the table.

Also, this distinction is not only dependent on the inherent verb semantics, but also affected by the nature of the construction, as in sentences like Example (104).

- (104) John smiled his way into the room.

9.2.3 Syntactic Role of the PP. The PP may serve as either **argument** or **adjunct**. This is a theoretical distinction, but we have argued that put-verbs serve as good candidates for verbs taking PP complements. We have also argued that PP arguments may be analyzed as LPs, and that the complement of put-verbs are best interpreted as locations. Put-verbs allow static and cofinal locatives, but not coinital locatives. Thus, we follow Kracht (2002) in that locatives are generally adjuncts, but may be selected as LPs under certain circumstances. This is shown in Examples (t.8) and (t.9), where the complement PP is analyzed as an LP, and where both static and directional (cofinal) locatives are accepted.

- (t.8) John la musa på_{st/cf} bordet.
John laid the mouse on/onto the table.
- (t.9) *John la musa fra vinduet.
*John laid the mouse from the window.

Kracht: Modificational PPs are adjuncts. Only when the PP is a mandatory argument, or the selecting head enters a relationship with the LP or NP level of the PP, is the PP regarded an argument.

ERG: A locative PP can be adjoined to any type of VP. In addition, directional verbs take directional PPs as arguments. PP arguments of put-verbs may be either directional or static, as described in Section 8.5. No treatment of location denoting arguments.

¹⁴ For details about this aspect we refer to the original article (Kracht 2002, pages 219–223).

9.2.4 *PP Transitivity*. Norwegian locative prepositions may **select PPs** (Example (t.10)), **NPs** (Example (t.11)), **no complement** (Example (t.12)) or **alternate** between no complement and PP selection (Example (t.13)). As the alternating prepositions take optional complements, they are ambiguous with respect to being the head of a complex PP or the first of two separate PPs (Example (t.13)). Furthermore, *under* in Example (t.13) is three-ways ambiguous with respect to mode (cf. Section 4.2). We also include an alternating locative (Example (t.14)) where the succeeding PP is blocked from being an LP complement, because the latter PP is cinitial mode, whereas alternating locatives only may select static PP complements as LPs.

- (t.10) Musa kom fra under bordet.
The mouse appeared from under the table.
- (t.11) Musa løp under bordet.
The mouse ran under the table.
- (t.12) Musa løp inn.
The mouse ran in.
- (t.13) Musa løp ned under bordet.
The mouse ran down under the table.
- (t.14) Musa løp ned fra bordet.
The mouse ran down from the table.

Kracht: Discusses PP and NP transitive locatives. PP transitive prepositions are treated as pure modalizers, where the selected PP is an LP. No treatment of intransitive and alternating locatives.

ERG: The analysis is discussed more in Section 8.5, where we gave examples of how ERG handled the different constructions. In particular, ERG does not distinguish between Examples (t.13) and (t.14) but ascribes the same number of analyses to them.

9.2.5 *Lexical Ambiguity*. Locatives exhibit **mode ambiguity**; recall Example (t.11), where *under* has three possible modes as discussed in Section 1 and at the end of Section 4.2. The mode ambiguity may be resolved by the syntactic context, as we see in Example (t.15). We underspecify the modes in the lexicon, and let the syntactic context disambiguate where possible.

- (t.15) Musa satt under bordet.
The mouse sat under the table.

Kracht: Does not treat lexical ambiguity directly.

ERG: Static/underspecified locatives are supertypes of directional locatives. Locatives are constrained to be directional when selected by a directional verb.

9.2.6 *Test Suite*. Based on these dimensions, we have constructed a small test suite in order to summarize and evaluate our approach. The test suite is given in Table 2. We present the total number of syntactic and semantic analyses produced by ERG. For our grammar we present the number of syntactic analyses and two numbers for

Table 2
Test suite.

Test Sentence	Our Analyses			Correct	ERG Analyses	
	Syntactic	Semantic (underspec.)	Semantic (total)	Semantic Analyses	Syntactic	Semantic
(t.1)	1	1	1	1	2	1
(t.2)	1	1	1	1	2	1
(t.3)	1	1	1	1	1	1
(t.4)	0	0	0	0	1	1
(t.5)	0	0	0	1	1	1
(t.6) ^a	1	1	1	1	1	1
(t.7) ^b	0	0	0	0	0	0
(t.8)	1	1	1	1	1	1
(t.9)	0	0	0	0	0	0
(t.10)	1	1	1	1	1	1
(t.11)	1	1	3	3 ^c	2	2
(t.12)	1	1	1	1	1	1
(t.13)	2	2	4	2	4	2
(t.14)	1	1	1	1	4	2
(t.15)	1	1	1	1	1	1

^a Only static readings are given in this row.

^b Only directional readings given in this row.

^c ERG does not distinguish between different directional modes. Given ERG's distinctions the correct number of analyses is 2.

the semantic analyses. The first is the number of different semantic representations produced by the system. They may be underspecified with respect to mode. The second number is the number of specified semantic analyses these underspecified analyses represent. We compare these to what we believe are the correct number of semantic analyses for each test sentence. We only include the analyses relevant to locatives from ERG, excluding analyses with non-locative interpretations of the PP, or where one of the prepositions in a complex locative PP is interpreted as a non-locative (e.g. analyses of *under* as a degree specifier in Examples (t.10) and (t.13)).

From the test suite, we see two problems for our analysis. The first is the undergeneration for Example (t.5), as we do not cope with static verbs and directional modification. This was discussed in Section 9.2.2. The second problem we find in our analysis of Example (t.13). For the reading where both locatives, *ned* and *under bordet*, are MPs, our analysis takes *ned* to be cofinal whereas *under bordet* is ascribed three different modalities. We do not find good arguments for claiming that this sentence has more readings than the cofinal+static reading besides the PP transitive reading where the two PPs are jointly cofinal. We have therefore indicated that the correct number of readings is 2 in this example. ERG gets a better result in this example because the directional reading of *under* is only obtained when *under* heads a PP which is argument to a directional verb. Because a directional verb only takes one PP argument, that reading is not obtainable in this sentence. On the other hand, ERG allows the PP *from the table* in Example (t.14) to either modify *down* or the VP. These two analyses result in two different semantic representations, which we find hard to motivate. Taking the correct semantic analyses as our goal we see that our recall is 14/15 and our precision is 14/16, whereas ERG achieves a recall of 14/14 and a precision of 14/16.

10. Conclusion

We have in this article described a treatment of locative prepositional phrases, which on the one hand is compatible with a compositional formal semantics and on the other hand is carried out in a framework for efficient computational grammars with an under-specified semantics. We believe the main strengths of our approach are the following:

1. **Uniformity:** Our analysis provides a uniform analysis of locatives, resulting in the same form on the semantic analyses independent of their syntactic realization.
2. **Simplicity:** We underspecify both mode ambiguity and the two roles of prepositions (as modificational or referential), by implementing the emptiness principle and claiming that any static preposition may serve as the LP complement of modalizers like *from*. Only modalizers need two lexical entries, all other locatives are encoded in a single lexical entry.
3. **Denotational Soundness:** All our representations have model-theoretic denotations. The interpretations are intuitively sound. We do not have any uninstantiated argument places.

Compared to other state-of-the-art approaches, we are able to give a far more detailed semantic analysis of the locative prepositions, and at the same time preserve a simple surface representation of the complex semantics of locative prepositions. This can be used in a translation system where it will distinguish between the mode aspect, which should be preserved from one language to the other, and the lexicalized locative expression, which may vary more unsystematically between the two languages.

Acknowledgments

We are grateful to our colleagues Stephan Oepen and Dan Flickinger for many enlightening discussions, for help with the LKB system and ERG, and for useful comments on the article. Needless to say, they do not agree with all we are saying and should not be held responsible for any mistakes. We are also grateful to the editors of the special issue and to the reviewers for their helpful advice.

References

- Alshawi, Hiyan, editor. 1992. *The Core Language Engine*. The MIT Press, Cambridge, MA.
- Bender, Emily M., Dan Flickinger, and Stephan Oepen. 2002. The grammar Matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In John Carroll, Nelleke Oostdijk, and Richard Sutcliffe, editors, *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, pages 8–14, Taiwan.
- Bierwisch, Manfred. 1988. On the grammar of local prepositions. In M. Bierwisch, W. Motsch, and I. Zimmermann, editors, *Syntax, Semantics und Lexicon*. Akademie Verlag, Berlin, pages 1–65.
- Bloom, Paul, Mary A. Peterson, Lynn Nadel, and Merrill F. Garrett, editors. 1996. *Language and Space*. The MIT Press, Cambridge, MA.
- Buschbeck-Wolf, Bianka and Rita Nübel. 1995. Die behandlung übersetzungsambiger präpositionen im transfer des verbmobil-demonstrators. Technical Report 87, Verbmobil, IBM, Heidelberg, Germany.
- Copestake, Ann. 2002. *Implementing Typed Feature Structure Grammars*. CSLI Publications, Stanford, CA.
- Copestake, Ann, Daniel Flickinger, Carl Pollard, and Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Journal of Research on Language and Computation*, 3(4):281–332.
- Copestake, Ann, Alex Lascarides, and Dan Flickinger. 2001. An algebra for semantic construction in constraint-based

- grammars. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL 2001)*, pages 132–139, Toulouse.
- Creary, Lewis G., J. Mark Gawron, and John Nerbonne. 1989. Reference to locations. In *Proceedings of the ACL 1989*, pages 42–50, Vancouver.
- Flickinger, Dan. 2002. On building a more efficient grammar by exploiting types. In Stephan Oepen, Dan Flickinger, Jun-ichi Tsujii, and Hans Uszkoreit, editors, *Collaborative Language Engineering. A Case Study in Efficient Grammar-Based Processing*, CSLI Lecture Notes. CSLI Publications, Stanford, CA, pages 1–17.
- Fong, Vivienne. 1997. *The Order of Things: What Directional Locatives Denote*. Ph.D. thesis, Stanford University.
- Herskovits, Annette. 1986. *Language and Spatial Cognition*. Cambridge University Press, Cambridge, England.
- Hjelmslev, Louis. 1935. La catégorie des cas. *Acta Jutlandica*, 7:1–184.
- Jackendoff, Ray. 1983. *Semantics and Cognition*. The MIT Press, Cambridge, MA.
- Jackendoff, Ray. 1990. *Semantic Structures*. The MIT Press, Cambridge, MA.
- Jørgensen, Fredrik. 2004. *The Semantic Representation of Locatives in Machine Translation*. Master's thesis, University of Oslo. Available at <http://wo.uio.no/as/Web0bjects/theses.woa/wa/these?WORKID=22120>.
- Kracht, Marcus. 2002. On the semantics of locatives. *Linguistics and Philosophy*, 25:157–232.
- Kracht, Markus. 2006. Directionality selection. In Patrick Saint-Dizier, editor, *Syntax and Semantics of Prepositions*, volume 29 of *Text, Speech and Language Technology*. Springer Netherlands, Dordrecht, pages 101–114.
- Levinson, S. C. 1996. Frames of reference and Molyneux's question: Cross linguistic evidence. In Paul Bloom, Mary A. Peterson, Lynn Nadel, and Merrill F. Garrett, editors, *Language and Space*. The MIT Press, Cambridge, MA, pages 109–169.
- Oepen, Stephan, Helge Dyvik, Jan Tore Lønning, Erik Velldal, Dorothee Beermann, John Carroll, Dan Flickinger, Lars Hellan, Janne Bondi Johannessen, Paul Meurer, Torbjørn Nordgård, and Victoria Rosén. 2004. Som å kapp-ete med trollet? Towards MRS-based Norwegian–English Machine Translation. In *Proceedings of the 10th International Conference on Theoretical and Methodological Issues in Machine Translation*, pages 11–20, Baltimore, MD.
- Pollard, Carl and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago, IL.
- Saint-Dizier, Patrick, editor. 2006. *Syntax and Semantics of Prepositions*, volume 29 of *Text, Speech and Language Technology*. Springer Netherlands, Dordrecht.
- Trujillo, Arturo. 1995. *Lexicalist Machine Translation of Spatial Prepositions*. Ph.D. thesis, University of Cambridge.
- van der Zee, Emile and Jon Slack, editors. 2003. *Representing Direction in Language and Space*. Oxford University Press, Oxford.
- Zwarts, Joost. 1997. Vectors as relative positions: A compositional semantics of modified PPs. *Journal of Semantics*, 14:57–86.
- Zwarts, Joost. 2003a. Paths round a prototype. In *ACL-SIGSEM Workshop: The Linguistic Dimensions of Prepositions and their Use in Computational Formalisms and Applications*, pages 228–238, Toulouse.
- Zwarts, Joost. 2003b. Vectors across spatial domains: From place to size, orientation, shape, and parts. In E. van der Zee and J. Slack, editors, *Representing Direction in Language and Space*. Oxford University Press, Oxford, pages 39–68.
- Zwarts, Joost. 2005. Prepositional aspect and the algebra of paths. *Linguistics and Philosophy*, 28(6):739–779.
- Zwarts, Joost and Yoad Winter. 2000. Vector space semantics: a model-theoretic analysis of locative prepositions. *Journal of Logic, Language and Information*, 9:169–211.