Combinatorial Aspects of PPs Headed by Raising Prepositions [1]

Beata Trawiński
Sonderforschungsbereich 441
Universität Tübingen
Nauklerstraße 35
72074 Tüebingen
trawinski@sfs.uni-tuebingen.de

Abstract
In this paper, semantic aspects of $P_1N_1P_2$ word sequences will be discussed. Based on syntactic analysis of [Trawiński 03], which assumes prepositions heading $P_1N_1P_2$NP combinations to be able to raise and realize syntactically complements of their arguments, we will investigate whether semantic representation of these expressions can be considered as an instance of the combinatoric semantics. We will investigate three German PPs involving expressions under consideration with respect to two criteria of internal semantic regularity adopted from [Sailer 00] and we will observe that the discussed expressions are not uniform with regard to the semantic properties. While the logical form of some of them can be computed by means of ordinary translations and a set of standard derivational operations, the other require additional handling methods. However, there are approaches available within the HPSG paradigm that are suited to account for these data. Here, we will briefly present the external selection approach of [Soehn 03] and the phrasal lexical entries approach of [Sailer 00] and we will show how they interact with the syntactic approach of [Trawiński 03].

Keywords
complex preposition, raising, idiosyncrasy, collocation, combinatoric semantics, bound word

1 Introduction

Regarding the complexity of the lexical material, constituting a preposition, a distinction in literature is commonly assumed between simple and complex prepositions. While simple prepositions consist of one word, complex prepositions are combinations of several lexical components. Among complex prepositions (CPs), sequences consisting of a preposition, a noun, and another preposition ($P_1N_1P_2$) are in many languages particularly frequent.[2] These word combinations are commonly considered to be unpredictable with regard to standard grammar regularities, and thus, often requiring additional description mechanisms. However, in [Trawiński 03] it has been shown that the syntax of $P_1N_1P_2$s can be described within the HPSG paradigm in the tradition of [Pollard and Sag 94] using a well-established in that grammar system mechanism of raising. Based on this approach, we will attempt in this paper to contribute to the treatment of $P_1N_1P_2$ expressions in describing semantic aspects of their grammar.

2 Syntactic Aspects

2.1 Some Empirical Observations

As $P_1N_1P_2$ expressions, we consider word combinations such as those in (1).

(1) an Hand von ('by means of'), in Zusammenhang mit ('in connection with'), unter Aufsicht von ('under survey of') ...[3]

[1] I would like to thank Manfred Sailer and Jan-Philipp Soehn for many helpful suggestions on the issue presented in this paper and Carmella Payne for help with English.

[2] Cf. [Lindqvist 94] , [Quirk and Mulholland 64], [Beneš 74], etc.

[3] The above collection is in no-way exhaustive. Moreover, it is unclear how extensive, at least approxi-

At first view, the interdependence between the particular elements of those expressions seems to defy standard constraints on the PP structure of German, since using PPs involving $P_1N_1P_2$ sequences such as *in Verbindung mit* ('in connection with') in the contexts exemplified in (2), we can observe many differences with the traditional PPs.

(2) In Verbindung mit diesem Problem möchte ich darauf hinweisen, dass ...
in connection with this problem would_like I DA_on point_out that
'In connection with this problem, I would like to point out that ...'

First of all, the noun *Verbindung* cannot be determined or quantified, nor can it be combined with possessive pronouns or prenominal genitives (cf. (3a)). Secondly, it cannot be modified (cf. (3b)). Finally, the PP *mit den Beratern* ('with the advisers') cannot be deleted (cf. (3c)).

(3) a. in *einer/ *der/ *seiner/ *Peters Verbindung mit diesem Problem ...
in a/ the/ his/ Peter's connection with this problem

b. in *enger/ *unerwarteter [Verbindung mit diesem Problem] *von dieser Woche/ *die uns betrifft, möchten wir ...
in close/ unexpected [connection with this problem] from this week/ which us concerns would_like we

c. * in Verbindung möchten wir ...
in connection would_like we

2.2 The Analysis of [Fries 88]

Based on empirical observations discussed above, the assumption can be made that the string *in Verbindung mit* ('in connection with') is a complex lexical sign. Thus, [Fries 88] provides for those PPs the structure in Figure 1.

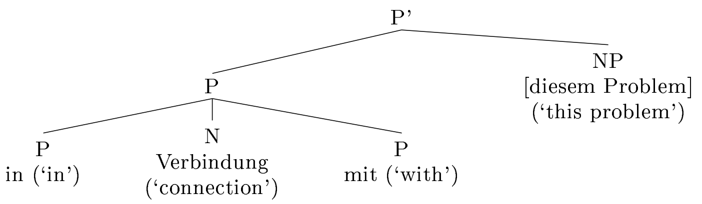

Figure 1: The structure of the PP *in Verbindung mit diesem Problem* ('in connection with this problem') in Fries (1988)

The preposition heading the entire phrase is a projection of three lexical categories which form together a complex lexical category, in this case, a preposition *in Verbindung mit*. This complex preposition then selects an NP forming a prepositional phrase.

The main problem with this analysis consists in the assumption that the preposition *mit* ('with') belongs to the complex preposition and cannot form a constituent with the NP *diesem Problem*. However, there are several data demonstrating the opposite.

Firstly, the combinations P_2NP where P_2 is realized by *von* ('of') can be replaced by the genitive; this replacement of *von* adheres to the restrictions on distribution of postnominal

mately, it is. For German, [Schröder 86] specifies more than 90 $P_1N_1P_2$ expressions, while e.g. [Beneš 74] itemizes 160 examples, thereby emphasizing the incompleteness of his list. In any case, these word combinations do not form a marginal class of expressions in contemporary German. For discussion on CPs of the discussed type in German see also [Maibauer 95].

genitives and *von*-PPs in German (cf. (4a)). Secondly, the discussed sequences can be substituted by *wo/da* expressions as in (4b), which are usually considered as proforms for PPs.

(4) a. mit Hilfe ??von dem Buch/ des Buches
with help of the book/ the book$_{GEN}$
'by dint of the book'

b. in Verbindung womit/damit
in connection WO_with/DA_with
'in connection with what/with it'

These observations imply that the P_2NP sequences form a constituent.

2.3 The Analysis of [Trawiński 03]

Considering all previous observations, one can assume that P_2NPs within $P_1N_1P_2NP$ expressions are lexically selected by N_1s, but they are realized syntactically by P_1s.

Based on this assumption, [Trawiński 03] provides for these expressions an analysis using the raising mechanism. This analysis assumes two uses of prepositions: the raising and the non-raising usages. The preposition *in* in (5a) occurs in a non-raising context, while the preposition *in* in (5b) occurs in a raising context.

(5) a. in einer engen Verbindung mit den Beratern
in a close connection with the advisors
'in a close connection with the advisors'

b. in Verbindung mit diesem Problem
in connection with this problem
'in connection with this problem'

The assumption is that both strings *mit diesem Problem* in (5b) and *mit den Beratern* in (5a) act as arguments of the noun *Verbindung* in that their syntactic properties are determined by this noun. Both *mit diesem Problem* and *mit den Beratern* are expected to be selected by *Verbindung* syntactically. Thus, in both cases, we proceed according to the standard methods of handling relational nouns which select prepositional arguments. This explains why the PP *mit diesem Problem* shares grammatical properties with the PP *mit den Beratern* and other ordinary PPs.

Furthermore, it will be assumed that in opposition to *in* in (5a), which subcategorizes the saturated NP, the preposition *in* in (5b) selects first the noun *Verbindung* (which does not realize its complement) and then selects the complement of *Verbindung*: the PP *mit diesem Problem*. That is, through an appropriate lexical principle of grammar specifying the valence of prepositions, the complement of the noun *Verbindung* is raised by *in* to become the complement of *in*, and to be realized by *in* syntactically.

In following, we will sketch the HPSG formalization of those assumptions.

To avoid redundancies in the lexicon, only one lexical entry for *in* will be specified, bearing partially underspecified information about its argument (cf. the relevant part of the lexical entry of the preposition *in* in AVM notation in Figure 2).
The only information about potential arguments of *in* provided in this lexical entry is that *in* can take only one argument, and this argument has to be a noun. Here, information about valence requirements of neither the selected noun, nor the preposition *in* are specified.

The syntactic selection properties of *in* are licensed by a constraint on the mapping of the elements of the ARG-ST list to the valence lists (cf. Figure 3).

$$\begin{bmatrix} \textit{word} \\ \text{PHON } \langle \textit{ in } \rangle \\ \text{ARG-ST } \left\langle \left[\text{LOC} \mid \text{CAT} \mid \text{HEAD } \textit{noun}\right] \right\rangle \\ \text{SYNS} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD } \textit{prep} \end{bmatrix}$$

Figure 2: The relevant part of the lexical entry of the preposition *in* ('in')

$$\forall \boxed{1} \forall \boxed{2} \left(\begin{bmatrix} \textit{word} \\ \text{SYNS} \mid \text{LOC} \mid \text{CAT} \begin{bmatrix} \text{HEAD } \textit{prep} \\ \text{ARG-ST } \boxed{1} \left\langle \left[\text{LOC} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS } \boxed{2}\right]\right\rangle \end{bmatrix} \end{bmatrix} \longrightarrow \left(\begin{array}{l} \boxed{1} = \left(\begin{array}{l} \left\langle \left[\text{LOC} \mid \text{CAT} \mid \text{VAL} \begin{bmatrix} \text{SPR } \langle\,\rangle \\ \text{SUBJ } \langle\,\rangle \\ \text{COMPS } \langle\,\rangle \end{bmatrix} \right] \right\rangle \vee \\ \left\langle \begin{bmatrix} \text{LEX } + \\ \text{LOC} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS } \langle \textit{synsem} \rangle \end{bmatrix} \right\rangle \end{array} \right) \wedge \\ \left[\text{SYNS} \mid \text{LOC} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS } \boxed{2} \otimes \boxed{1} \right] \end{array} \right) \right)$$

Figure 3: ARG-ST Mapping Lexical Principle for Prepositions

In order to facilitate prepositions to subcategorize nouns with an unsaturated complement, and then also to select the complements of those nouns, the list of complements which are syntacticly selected by a preposition has been specified as a concatenation of its own ARG-ST list and the list of complements of its argument.[4]

It should be mentioned that the raising of more than one nominal complement result in ungrammatical constructions like those in (6).

(6) in Verbindung *[der Regierung] mit diesem Problem ...
in connection the government$_{GEN}$ with this problem

To avoid this problem, the ARG-ST value of prepositions has been restricted to the lists containing either one saturated element, or to the lists containing one element with a singleton COMPS list. Additionally, the LEX value of the second disjunct has been specified to be + with the idea of marking objects that have realized none of their complements. This restriction rules out the selection of relational nouns that have already realized one of their complements (cf. 7).

(7) a. *in [Verbindung der Regierung] [mit diesem Problem] ...
in connection the government$_{GEN}$ with this problem

The structure in Figure 4 exemplifies the interaction of the above assumptions in licensing a PP headed by a raising preposition.
Due to the ARG-ST Mapping Lexical Principle for Prepositions in Figure 3, a preposition *in* can be licensed, which takes one nominal argument with one unrealized complement.

[4] We assume, as [Meurers 97] does, that argument raising takes place only with respect to the valence attributes, not to the ARG-ST LIST. The intuition behind this has to do with the character of the ARG-ST list as the direct syntactic reflexion of the semantic argument structure.

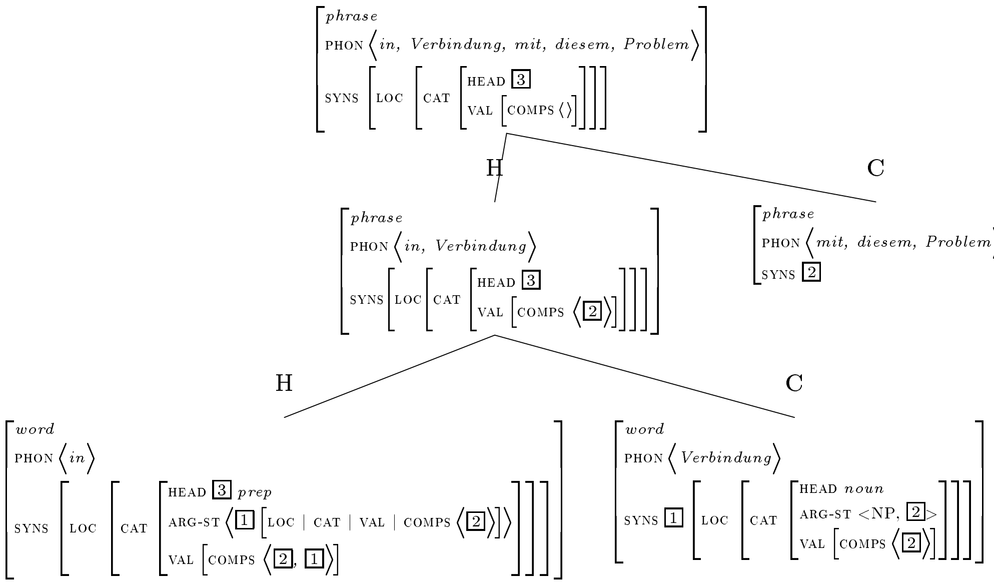

Figure 4: The structure of the PP *in Verbindung mit diesem Problem*

Thus the syntactic and semantic properties of that complement are determined not by the preposition but by the noun. Both the noun and its unrealized complement are mapped to the COMPS list of *in* and, according to the constraints on the head-complement-structures for prepositions, they are syntactically selected by *in*.

The first complement that *in* selects is the noun. By virtue of selectional requirements of restrictive adjectives as well as prepositions modifying nouns, that are specified as combining with complement-saturated nouns only, the modifying of complement-unsaturated nouns is blocked. The same restriction holds for determiners and quantifiers in German. These constraints, existing in the grammar independently of the principles of the CPs syntax, explain the apparent lexical fixedness of the P_1N_1 sequences (cf. (3a) and (3b)) without additional stipulations.

Further on, the preposition *in* selects the complement of the noun as its own complement, forming a PP.

Exactly the same lexical entry for preposition *in* and the same set of principles license PPs headed by non-raising prepositions such as the PP *in einer engen Verbindung mit den Beratern* ('in close connection with the advisers').

3 Semantic Aspects

In the previous section, we have shown following [Trawiński 03] that the syntactic structure of PPs consisting of $P_1N_1P_2NP$ sequences can be described by use of the raising mechanism that enables prepositions to raise and realize syntactically complements of their arguments.[5] We can thus conclude that these expressions are licensed by virtue of regular principles of syntax. It might therefore be expected that also the meaning of that PPs is an instance of likewise regular compositional semantics. However, the discussion below will show that this is not always the case. In fact, PPs of the discussed type are not uniform concerning their semantic properties. To illustrate this, we will investigate three PPs corresponding to $P_1N_1P_2NP$ patterns with respect to the following two aspects referred to as

[5]It should be mentioned that the proposed analysis entails a technique which is already well established in the HPSG-based studies. On further applications of the raising mechanism see e.g. [Hinrichs and Nakazawa 89], [Hinrichs and Nakazawa 94], [Meurers 00] or [De Kuthy 00].

Regularity Properties (RPs)[6]:

RP 1: Every element in the PP can be attributed some meaning with which it occurs also outside the particular combination under consideration.

RP 2: The meaning of the entire PP is arrived at by combining the meanings of its parts in a regular way.

If at least one of those criteria is violated by a given $P_1N_1P_2NP$ sequence, then this situation will be considered as an evidence for semantic irregularity (i.e. idiomatic character) of that expression, and will thus argue for assuming a special mechanism to account for the semantic representation of that PP.

3.1 The Meaning of Freely Combined PPs

First of all, we will investigate freely combined PPs with respect to the criteria in RP 1 and RP 2 in order to obtain a representing example of a non-idiomatic PP.[7] We will consider thus an ordinary PP *in einem Buch von Chomsky* ('in a book by Chomsky') as used in syntactic contexts such as those in (8).

(8)	Er	hat	über	diese	Theorie	*in*	*einem*	*Buch*	*von*	*Chomsky*	gelesen.
	he	has	about	this	theory	in	a	book	of	Chomsky	read

'He read about this theory in a book by Chomsky.'

In assigning every element of the above PP a semantic representation and based on the syntactic structure such as those in Figure 5, we can compute the meaning of the entire PP by combining the meanings of its particular elements in a way assumed in the common semantic theories, i.e. by functional application (cf. Figure 5).[8] As proposed in [Sailer 00], the terms occurring in the tree in Figure 5 instantiate the values of the HPSG CONTENT attribute of the particular signs.

We can thus see that the second criterion of regularity is met by this PP. Also, the first condition is satisfied, since every element of the discussed PP can appear with exactly the same meaning with which it appears in this particular PP in other syntactic contexts.

Based on the fact that the PP *in einem Buch von Chomsky* satisfies the both conditions on regularity, it can be considered as a semantically regular non-idiomatic expression and can thus be licensed by regular semantic constraints of grammar.

3.2 The Meaning of $P_1N_1P_2$ combinations

In this section, we will investigate three types of PPs corresponding to $P_1N_1P_2NP$ patterns that behave differently with respect to the conditions introduced as RP 1 and RP 2 as criteria of semantic regularity for syntactically complex expressions.

[6]These two criteria are based on the semantic criteria of regularity which are assumed by [Sailer 00] to account for idiomatic VPs. Here, we adopt them for PPs.

[7]This paper focuses on PPs headed by referential, i.e. semantically non-empty prepositions, that function as modifiers within syntactic structures. Prepositions heading PPs that act as complements are considered here as semantically vacuous items denoting the identity function of type $\lambda x.x$. Thus, the semantic representation of the PP *von Chomsky* ('by Chomsky') as used in (8) is assumed to have a form of a constant *chomsky*.

[8]The noun *Buch* as it occurs in (8) is assumed to be a relational noun, i.e. in the expression **buch'**$(y)(x)$, x is interpreted as the book and y as its author.

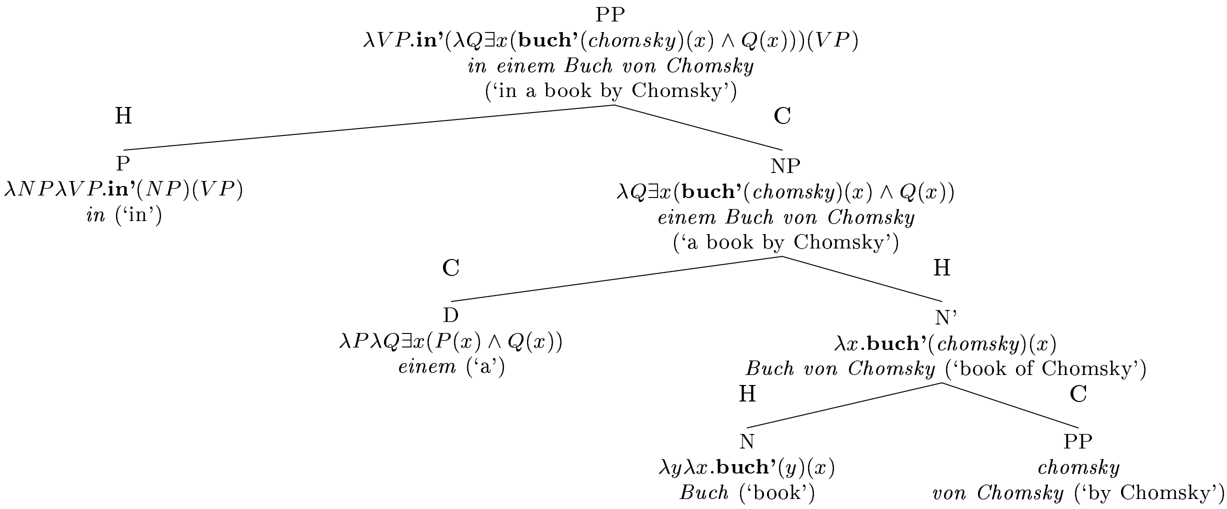

Figure 5: The structure of the PP *in einem Buch von Chomsky* ('in a book by Chomsky')

First of all, we will examine the PP *in Verbindung mit x* ('in connection with x') with respect to the condition in RP 1, according to that every element of that PP can be assigned some meaning whose occurrence with that element is no way limited to this particular PP. It is obvious that the noun *Verbindung* in the PP *in Verbindung mit x* refers to a relation which contains tripples consisting of a connection or contact event, an individual y, and an individual x as it does in other syntactic contexts (cf. (9)).

(9) Eine Verbindung von bin Laden mit Sadam Husain wurde nie nachgewisen.
a connection of bin Laden with Sadam Husain was never proven
'A connection of bin Laden with Sadam Husain has never been proven'.

For the noun *Verbindung* in (9), we can thus assume a denotation of the form $\lambda y\lambda x\lambda e.\mathbf{verbindung'}(e)(y)(x)$. However, the value of y does not necessarily have to be instantiated (cf. (10)). In such cases, it is plausible to assume that the semantic representation of the noun *Verbindung* as it occurs in (10) is a term such as $\lambda x\lambda e.\exists y\mathbf{verbindung'}(e)(y)(x)$.

(10) Eine Verbindung PRO mit Sadam Husain ist zurzeit ausgeschlossen.
a connection PRO mit Sadam Husain is currently impossible
'A connection with Sadam Husain is currently not possible.'

Considering the fact that bivalent nouns occurring as objects of raising prepositions in contexts such as *in Verbindung mit x* may not realize their subjects (cf. (6)), we can plausibly assume the second term to be a translation of *Verbindung* when combining with the raising preposition *in*. However, one variable more within that translation has to be abandoned, i.e. the event variable e. Since no proper quantification over the noun *Verbindung* is possible in the PP *in Verbindung mit x*, we assume a default existential quantifier that binds the event variable e, i.e. $\lambda x.\exists e\exists y\mathbf{verbindung'}(e)(y)(x)$. This assumption would partially explain why no number contrasts are possible in combinations under consideration. Apparently, the plural operator introduced by plural forms seems not to be compatible with the default existential quantifier introduced by the translation of *Verbindung*.[9] In any case, we assume that the meaning of the noun *Verbindung* in the PP *in Verbindung mit x* is derived from its base translation, that is appropriate for this noun in combinations other than the discussed PP as well.

[9]However, the problem of pluralization of deverbal N_1s denoting events mainly goes back to the issues related to the nominalization, which will not be discussed in this paper.

Similarly, the preposition *in* occurs in the PP *in Verbindung mit x* with its metaphorical non-spatial meaning, with which it appears also in many other combinations (cf. (11)). For the preposition *in* acting as a VP modifier, we can thus assume a translation of the form $\lambda P \lambda Q.\mathbf{in'}(P)(Q)$.

(11) in einer Beziehung/ einer Relation/ einem Kontakt sein
in a connection/ a relationship/ a contact be
'to be in a connection/ a relationship/ a contact'

Also, the occurrence of the preposition *mit* with its meaning of the identity function of the form $\lambda x.x$ is not limited to the *in Verbindung mit x* combination (cf. (12)).

(12) sich mit Semantik beschäftigen/ mit Fisch handeln
self with semantics deal/ with fish deal
'to deal with semantics/ to deal in fish'

Based on this, we can conclude that the PP *in Verbindung mit x* satisfies the condition in RP 1. We assume also that the second criterion of regularity is satisfied by that PP. Using certain type shifting operations in style of *argument raising* or *value raising* introduced within a semantic framework of *Flexible Montague Grammar* of [Hendriks 93] or in style of the *Adjunct Introduction Derivational Rule* of [Richter and Sailer 99] to account for negation in French, we can derive the meaning of the entire PP in a combinatoric way.

One can certainly treat most $P_1N_1P_2NP$s in which N_1s are deverbal event nominalizations as semantically regular, i.e. licensed by regular translations and regular derivational operations. The assumption that they are both syntactically and semantically freely combined explains their high productivity in contemporary German.

However, not every type of $P_1N_1P_2NP$ expressions can be handled in that way. Thus, $P_1N_1P_2NP$s of type *an Hand von x* ('by means of x'), *an Stelle von x* ('in lieu of x') or *auf Grund von x* ('by virtue of x') are significantly less frequent in German. They consist of lexical entities every one of which appears also outside a particular PP, bearing however a complete different meaning. Moreover, considering the meanings of entire PPs of the discussed type, it is highly problematic to assign any meanings to the particular elements of those PPs such that they could be combined into the meaning of the entire PP. Therefore, neither of RPs are satisfied by this class of $P_1N_1P_2NP$s.

Another case of PPs involving $P_1N_1P_2$ sequences form combinations that contain a so-called *bound word*, e.g. *in Anbetracht von x* ('in consideration of x'). Here, we can consider the entire PP as semantically decomposable, and, thus satisfying the condition of semantical regularity in RP 2. However, the first condition is violated, since not all components of that PP occur with the meaning that they have inside of that PP also in other contexts. The distribution of the noun *Anbetracht* is limited in German exclusively to the combination with the preposition *in*.[10]

Thus, to account for semantic representation of PPs such as *an Hand von x* and *in Anbetracht von x*, we need some special mechanisms.

3.3 Available HPSG Analyses

In Section 3.2, we have identified three classes of PPs involving $P_1N_1P_2$ expressions with respect to the conditions of semantic regularity RP 1 and RP 2. We have seen that the

[10] For the treatment of *bound words* acting as prepositional objects within the HPSG grammar framework see [Soehn 03] and [Soehn and Sailer 03].

expressions under consideration do not show uniform properties with regard to semantic aspects. While the semantic representation of PPs whose N_1 is realized by a deverbal event nominalization can be licensed by ordinary translation rules and a set of common derivational operations in the way shown in Figure 5 for freely combined PPs, the computation of a semantic representation of PPs containing *bound words* and PPs of type *an Hand von* is more challenging.[11]

Nevertheless, there are HPSG-based approaches that introduce necessary means to account for such phenomena, the external selection approach of [Soehn 03] and the phrasal lexical entry approach of [Sailer 00]. In the following, we will sketch the main ideas of both theories and we will show how they can be used to account for the more idiosyncratic PPs.

The external selection approach of [Soehn 03] generalizes the available HPSG external selection mechanisms (cf. the MOD and SPEC features) in that it assumes that in **every** type of phrase, the non-head daughter is permitted to determine syntactic and semantic properties of the head daughter. This idea is realized by the attribute XSEL (*external selection*), which is assumed to be appropriate for the sort *head* and to take a *synsem* object as its value. A XSEL value of a *bound word* such *Anbetracht* will thus be the SYNSEM value of the preposition *in*. This lexical constraint will ensure the occurrence of *Anbetracht* exclusively within a PP headed by the preposition *in*. For freely occurring words, the value of their XSEL feature will be underspecified.

The PP *in Anbetracht von x* is a $P_1N_1P_2NP$ expression, however. That is, the preposition *in* raises and realizes syntacticly the argument of the noun *Anbetracht*, i.e. the PP *von x*. The lexical entry for the noun *Anbetracht* in Figure 6 quoted from Soehn, p.c., shows that we can merge the external selection approach of [Soehn 03] with the complement raising approach smoothly.

$$\left[\begin{array}{l} \textit{word} \\ \text{PHON } \langle \textit{ Anbetracht } \rangle \\ \text{SYNS } \boxed{1} \left[\text{LOC} \left[\text{CAT} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{noun} \\ \text{CASE } \textit{dat} \\ \text{XSEL} \left[\text{LOC} \left[\text{CAT} \left[\begin{array}{l} \text{HEAD } \textit{prep} \\ \text{VAL} \mid \text{COMPS } \langle \boxed{2}, \boxed{1} \rangle \\ \text{ARG-ST } \langle \boxed{1} \left[\text{LOC} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS } \langle \boxed{2} \rangle \right] \rangle \end{array}\right]\right]\right] \end{array}\right] \\ \text{VAL} \mid \text{COMPS } \langle \boxed{2} \rangle \\ \text{ARG-ST } \langle \boxed{2} \rangle \end{array}\right]\right]\right] \end{array}\right]$$

Figure 6: The relevant part of the lexical entry of the noun *Anbetracht*

If we assume the usual non-spatial meaning for the preposition *in* selecting *Anbetracht*, then we can derive the meaning of the entire PP in a fully regular compositional way.

However, a compositional analysis will not be possible in the case of $P_1N_1P_2NP$ expressions of type *an Hand von x*. As we have ascertained above, the meaning of the entire PP cannot be obtained from the meanings of their parts. Only in that particular combination the sequence *an Hand von x* means what it means. Therefore, the assumption seems to be plausible that as well as words, these expressions are licensed directly by the lexicon.

This assumption underlies the HPSG-based approach of phrasal lexical entries of [Sailer 00]. According to this approach, syntactically complex expressions (*phrases*) which show some

[11] The problem related to the variation of the degree of semantic as well as syntactic interdependence between particular elements of syntactically complex expressions considered generally as multiword units is also mentioned in other approaches, such as [Copestake et al 02] and [Sag et al 02].

internal irregularities are encoded directly in the lexicon as phrasal lexical signs and can be identified and distinguished from internally regular syntactically complex expressions by virtue of an attribute taking boolean values. Thus, the attribute COLL (for *Context of Lexical Licensing*) will be declared appropriate for the sort *phrase* that takes + or − as its value. Thereby all regular phrases have a COLL value of sort −, whereas internally irregular complex expressions are specified to have a COLL value of sort +.[12]

According to this approach, PPs such as *an Hand von x* can be licensed to introduce a meaning which does not relate to the semantic contribution of their parts, but is only associated with that particular word combination. Figure 7 illustrates the derivation of the PP *an Hand von Prolog* with the application of the simplified COLL mechanism.

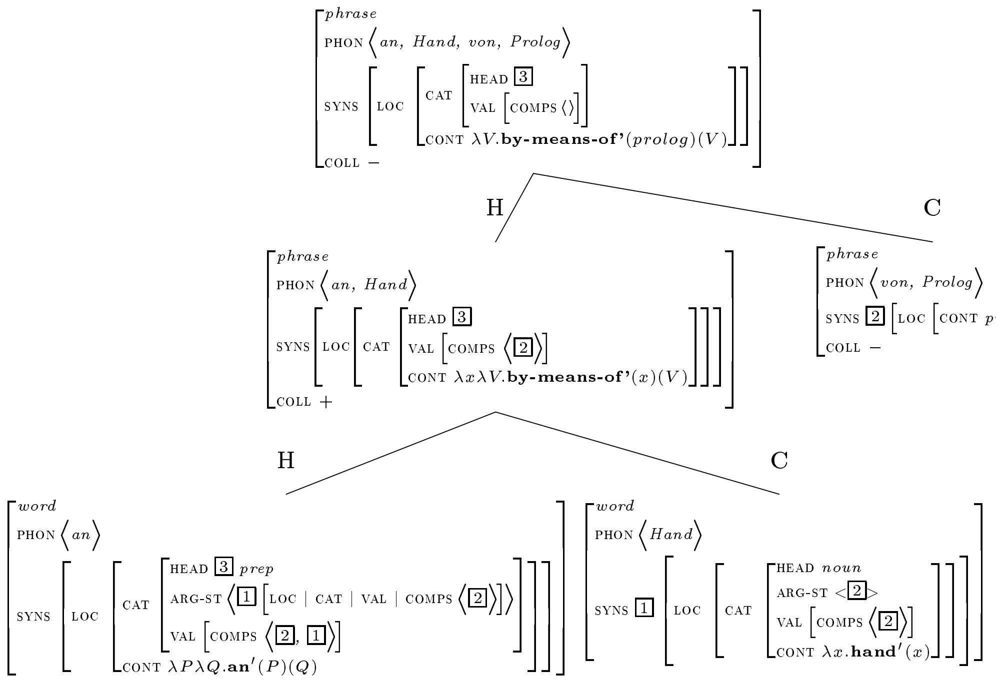

Figure 7: The structure of the PP *an Hand von Prolog* ('by means of Prolog')

As we can see in Figure 7, the COLL value of phrases *von Prolog* and *an Hand von Prolog* is specified as −, since these entities are expressions licensed by regular constraints of grammar. In contrast, the COLL value of the phrase *an Hand* is specified as +. As an internally irregular expression, the phrase *an Hand* is licensed immediately by the lexicon providing a new semantic constant.

In principle, if we use the full phrasal lexical entries approach using the COLL mechanism of [Sailer 00], we can describe any idiosyncratic phenomenon, including those related to *bound words*.[13] Thus, PPs such *in Anbetracht von x* discussed above can be captured by that theory as well. Nevertheless, it would be interesting to investigate precisely what potentialities the interaction of the XSEL and the COLL mechanisms could provide for describing

[12] To account for phenomena discussed in this paper, it is sufficient to adopt the simplified variant of the usage of the COLL attribute presented in Section 8.1 of [Sailer 00]. In Section 8.3 of [Sailer 00], it will be assumed that the COLL attribute takes list of signs as its value, such that for every non-lexical sign, the value of the COLL feature is specified to be an empty list and the value of every lexical sign is a singleton list containing an element that is identical with the root sign of the utterance in which this lexical sign occurs. By virtue of that mechanism any distributional phenomena can be described. However, for our purposes, that version of the approach in [Sailer 00] seems too powerful.

[13] For treatment of *bound words* within the HPSG grammar system using the COLL tool, see [Richter and Sailer 02]. See also [Sailer and Richter 02] on account for collocational requirements of some German verbs using the same technique.

linguistic phenomena.

4 Summary

[Trawiński 03] discusses syntactic properties of PPs involving $P_1N_1P_2$ sequences which are the basis for the complement raising analysis. Based on this analysis, we have investigated further properties of $P_1N_1P_2$s focusing on semantic aspects. The objective of our investigations was to examine whether the semantic representation of these expressions can be derived combinatorial. We have thereby seen that the discussed expressions are not uniform with regard to their semantic behavior, forming three classes: $P_1N_1P_2$s that can be analyzed compositionally (*in Verbindung mit*), $P_1N_1P_2$s involving *bound words* that can be treated within the combinatorial semantics as well, but requiring some mechanism which describes distributional properties of the particular *bound words* (*in Anbetracht von*), and $P_1N_1P_2$s that cannot be handled by virtue of common derivational methods (*an Hand von*). Nevertheless, we have seen that there are HPSG-based approaches that account for all that data. Possibly, merging the syntactic analysis for $P_1N_1P_2$ sequences of [Trawiński 03] with the external selection approach of [Soehn 03] and the phrasal lexical entry approach of [Sailer 00] could provide a uniform and satisfactory theory of $P_1N_1P_2$ combinations.

References

Beneš, E. (1974). Präpositionswertige Präpositionalfügungen. In U. Engel and P. Grebe (Eds.), *Sprachsystem und Sprachgebrauch. Festschrift für Hugo Moser zum 65. Geburtstag. Teil I.*, Number 33 in Sprache der Gegenwart, pp. 33–52. Düsseldorf: Schwann.

Copestake, A., F. Lambeau, A. Villavicencio, F. Bond, T. Baldwin, I. A. Sag, and D. Flickinger (2002). Multiword Expressions: Linguistic Precision and Reusability. In *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002)*, Las Palmas, Canary Islands.

De Kuthy, K. (2000). *Discontinuous NPs in German — A Case Study of the Interaction of Syntax, Semantics and Pragmatics.* Ph. D. thesis, Universität des Saarlandes, Saarbrücken.

Fries, N. (1988). *Präpositionen und Präpositionalphrasen im Deutschen und im Neugriechischen.* Number 208 in Linguistische Arbeiten. Tübingen: Max Niemeyer Verlag.

Hendriks, H. (1993). *Studied Flexibility.* ILLC Dissertation Series 1995-5. Institute for Logic, Language and Computation, Amsterdam.

Hinrichs, E. and T. Nakazawa (1989). Flipped Out: Aux in German. In *Papers from the 25th Regional Meeting of the CLS*, Chicago, Illinois, pp. 193–202.

Hinrichs, E. and T. Nakazawa (1994). Linearizing AUXs in German Verbal Complexes. In J. Nerbonne, K. Netter, and C. Pollard (Eds.), *German in Head-Driven Phrase Structure Grammar*, Number 46 in CSLI Lecture Notes, pp. 11–37. Stanford, CA: CSLI Publications.

Lindqvist, C. (1994). *Zur Entstehung der Präpositionen im Deutschen und Schwedischen.* Tübingen: Max Niemeyer Verlag.

Meibauer, J. (1995). Komplexe Präpositionen — Grammatikalisierung, Metapher, Implikatur und *Division of Pragmatic Labour*. In F. Liedtke (Ed.), *Implikaturen. Grammatische und pragmatische Analysen*, Number 343 in Linguistische Arbeiten, pp. 67–74. Tübingen: Max Niemeyer Verlag.

Meurers, W. D. (1997). Using Lexical Principles in HPSG to Generalize over Valence Properties. In *Proceedings of the Third Conference on Formal Grammar*, Aix-en-Provence, France.

Meurers, W. D. (2000). *Lexical Generalizations in the Syntax of German Non-Finite Constructions*. Ph. D. thesis, Eberhard-Karls-Universität Tübingen. Published as: Arbeitspapiere des SFB 340, Nr. 145.

Pollard, C. J. and I. A. Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago: The University of Chicago Press.

Quirk, R. and J. Mulholland (1964). Complex Prepositions and Related Sequences. In *English studies presented to R. W. Zandvoord on the occasion of the 70th birthday, Supplement to Vol. 45*, Amsterdam, pp. 64–73.

Richter, F. and M. Sailer (1999). A Lexicalist Collocation Analysis of Sentential Negation and Negative Concord in French. In V. Kordoni (Ed.), *Tübingen Studies in Head-Driven Phrase Structure Grammar*, Arbeitspapiere des SFB 340, Nr. 132, Volume 1, pp. 231–300. Universität Tübingen.

Richter, F. and M. Sailer (2002). Cranberry Words in Formal Grammar. In C. Beyssade, O. Bonami, P. Cabredo Hofherr, and F. Corblin (Eds.), *Empirical Issues in Formal Syntax and Semantics*, Volume 4. Presses Universitaires de Paris-Sorbonne, Paris.

Sag, I. A., T. Baldwin, F. Bond, A. Copestake, and D. Flickinger (2002). Multiword Expressions: A Pain in the Neck for NLP. In A. Gelbukh (Ed.), *Computational Linguistics and Intelligent Text Processing: Third International Conference: CICLing-2002*, Hiedelberg, Berlin. Springer-Verlag.

Sailer, M. (Version of June 29th, 2000). *Combinatorial Semantics and Idiomatic Expressions in Head-Driven Phrase Structure Grammar*. Ph. D. thesis, Eberhard-Karls-Universität Tübingen, Tübingen.

Sailer, M. and F. Richter (2002). Not for Love or Money: Collocations! In G. Jäger, P. Monachesi, G. Penn, and S. Wintner (Eds.), *Proceedings of Formal Grammar 2002*, pp. 149–160.

Schröder, J. (1986). *Lexikon deutscher Präpositionen*. Leipzig: Verlag Enzyklopädie.

Soehn, J.-P. (2003). Von Geisterhand zu Potte gekommen. Eine HPSG-Analyse von PPs mit unikaler Komponente. Master's thesis, Universität Tübingen, Seminar für Sprachwissenschaft.

Soehn, J.-P. and M. Sailer (2003). At First Blush on Tenterhooks. About Selectional Restrictions Imposed by Nonheads. In G. Jäger, P. Monachesi, G. Penn, and S. Wintner (Eds.), *Proceedings of Formal Grammar 2003*, pp. 149–161.

Trawiński, B. (2003). Licensing Complex Prepositions via Lexical Constraints. In F. Bond, A. Korhonen, D. McCarthy, and A. Villavicencio (Eds.), *Proceedings of the ACL 2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, Sapporo, Japan, pp. 97–104.