Chapter 1

COMBINATORIAL ASPECTS OF COLLOCATIONAL PREPOSITIONAL PHRASES

Beata Trawiński
University of Tübingen, Collaborative Research Centre 441
Nauklerstraße 35, 72074 Tübingen, Germany
trawinski@sfs.uni-tuebingen.de

Manfred Sailer
University of Göttingen, Department of English Studies
Käte-Hamburger-Weg 3, 37073 Göttingen, Germany
manfred.sailer@phil.uni-goettingen.de

Jan-Philipp Soehn
University of Tübingen, Collaborative Research Centre 441
Nauklerstraße 35, 72074 Tübingen, Germany
soehn@sfs.uni-tuebingen.de

Abstract In this paper we will discuss semantic aspects of collocational prepositional phrases (CPPs) consisting of $P_1 N_1 P_2 NP$ sequences. Based on the syntactic analysis in (Trawiński, 2003), which assumes prepositions heading $P_1 N_1 P_2 NP$ combinations to be able to raise and syntactically realize complements of their arguments, we will investigate whether the semantic representations of these expressions can be derived compositionally. We will discuss German CPPs with respect to two criteria of internal semantic regularity taken from (Sailer, 2003), and we will observe that the expressions in question are not uniform with regard to their semantic properties. While the logical form of some of them can be computed by means of ordinary meaning assignment and a set of standard derivational operations, others require additional handling methods. However, there are approaches available within the HPSG paradigm which are able to account for these data. Here we will briefly present the external selection approach of (Soehn, 2003) and the phrasal lexical entries approach of (Sailer, 2003), and we will demonstrate how they interact with the syntactic approach of (Trawiński, 2003).

Keywords: bound word, collocation, complex preposition, compositional semantics, HPSG, phrasal lexical entry, raising

Introduction

Among *collocational prepositional phrases* (CPPs), sequences consisting of a preposition, a noun, a second preposition, and an NP ($P_1N_1P_2NP$) occur particularly frequently in many languages.[1] These combinations are collocational in the sense of exhibiting a high degree of lexical fixedness. CPPs are commonly considered to be unpredictable with regard to standard grammar regularities. However, (Trawiński, 2003) has shown that the syntax of German CPPs can be described within HPSG (Pollard and Sag, 1994) using the well established mechanism of raising. Based on this syntactic approach, we will describe the semantic aspects of German CPPs. We will distinguish CPPs of different semantic regularity and combine independently motivated accounts to capture these expressions.

1. Syntactic Aspects

1.1 Some Empirical Observations

We consider the following word combinations to be $P_1N_1P_2$ expressions.

(1) an Hand von (at hand of, 'by means of'), in Verbindung mit (in connection with, 'in connection with'), unter Aufsicht von (under survey of, 'under the supervision of') ...[2]

At first glance, the interdependence between the particular elements of these expressions seems to defy standard constraints on the PP structure of German; on examining 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 compared to traditional PPs.

(2) In Verbindung mit diesem Problem will ich bemerken, dass ...
in connection with this problem want I note that
'In connection with this problem, I want to point out that ...'

First of all, the noun *Verbindung* ('connection') cannot combine with a determiner, a quantifier, a possessive pronoun or a prenominal genitive (cf. (3a)). Secondly, it cannot be modified (cf. (3b)). Finally, the PP *mit diesem Problem* ('with this problem') cannot be omitted (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]
in close/ unexpected [connection with this problem]
*von dieser Woche/ *die uns betrifft, will ich ...
from this week/ which us concerns want I

c. * in Verbindung will ich ...
in connection want I

Based on these observations, it is often assumed that the string *in Verbindung mit* ('in connection with') is a complex lexical sign (cf. the structure in (4) provided for those PPs by (Fries, 1988)).

(4) $[_{P'}[_{P}$ in$]$ $[_{N}$ Verbindung$]$ $[_{P}$ mit$]]$ $[_{NP}$ diesem Problem$]]$

The preposition heading the entire phrase is a projection of three lexical categories which together form a complex lexical category, in this case a preposition *in Verbindung mit* ('in connection with'). 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* ('this 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* ('of') adheres to the restrictions on distribution of postnominal genitives and *von*-PPs in German (cf. (5a)). Secondly, the sequences in question can be substituted by *wo-/da-* expressions as in (5b), which are usually considered as proforms for PPs. These observations imply that the P_2NP sequences form a constituent.

(5) a. an Hand von zwei Beispielen/ zweier Beispiele
by means of two examples/ two examples$_{GEN}$
'by means of two examples'

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

Taking all previous observations into consideration, one can conclude that within a $P_1N_1P_2NP$ expression the P_2NP is lexically selected by N_1, but realized as a syntactic sister of a P_1N_1 complex.

1.2 Raising Analysis

Based on the above generalization, (Trawiński, 2003) provides an analysis for these expressions using the raising mechanism.[3] We will outline here the HPSG formalization of this analysis.

To avoid redundancies in the lexicon, only one lexical entry for *in* ('in') will be specified (cf. Figure 1.1), bearing underspecified information about its argument's degree of saturation. The syntactic selection properties of *in* ('in') are licensed by a constraint on the mapping of the elements of the ARG-ST list to the valence lists (cf. Figure 1.2). In order to enable prepositions to subcategorize nouns with an unsaturated complement, and then also to select the complements of those nouns, the list of complements which are syntactically

selected by a preposition is specified as a concatenation of its own ARG-ST list and the COMPS list of its argument (cf. $\boxed{2} \otimes \boxed{1}$).

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

Figure 1.1. The relevant part of the lexical entry of the preposition *in* ('in')

$$\forall \boxed{1} \forall \boxed{2}$$

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

Figure 1.2. ARG-ST Mapping Lexical Principle for Prepositions

It should be mentioned that the raising of more than one nominal complement results in ungrammatical constructions such as the following:

(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 a preposition is restricted to be either a list with one saturated element, or a list containing one element with a singleton COMPS list (cf.$\boxed{1}$). Additionally, the LEX value of the second disjunct is specified as $+$. This marks objects which have not realized any of their complements. This restriction rules out the selection of relational nouns which have already realized one of their complements (cf. (7)).

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

The structure in Figure 1.3 exemplifies the interaction of our assumptions regarding the licensing of a PP headed by a raising preposition. According to the ARG-ST Mapping Lexical Principle for Prepositions in Figure 1.2 the preposition *in* ('in') can take one nominal argument with one unrealized complement. Thus the syntactic and semantic properties of this 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* ('in'), and, according to the constraints on the head-complement-structures for prepositions, they are syntactically selected by *in* ('in').

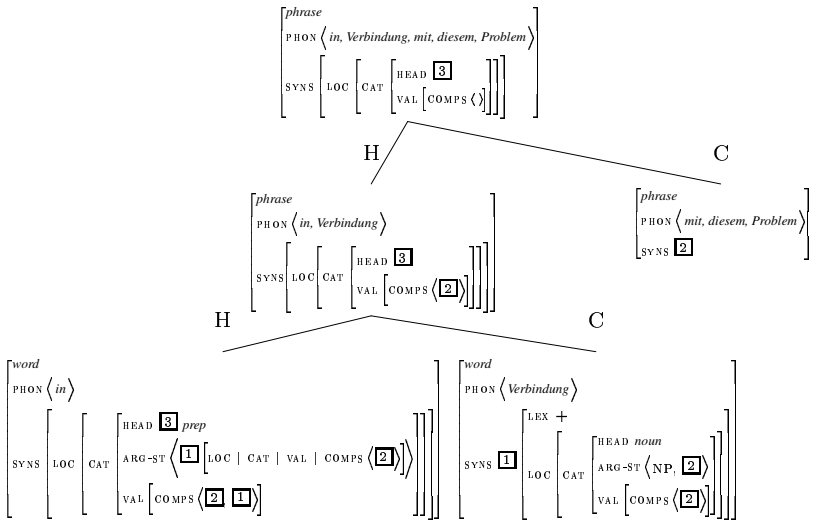

Figure 1.3. The structure of the PP *in Verbindung mit diesem Problem*

The empirical observations of Section 1.1 can be explained by this analysis. The first complement selected by *in* ('in') is the lexical noun. Restrictive adjectives or modifying PPs are both specified as combining only with complement-saturated nouns. Thus, adjunction to 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 CPP's syntax, explain the apparent lexical fixedness of the P_1N_1 sequences (cf. (3a) and (3b)) without additional stipulations. The combination *in Verbindung* ('in connection') selects the complement of the noun as its own complement, forming a PP.

Exactly the same lexical entry for *in* ('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 a close connection with the advisers').

2. Semantic Aspects

In the previous section we have argued that the syntactic structure of CPPs consisting of $P_1N_1P_2NP$ sequences can be described by use of the raising mechanism which enables prepositions to raise and syntactically realize complements of their arguments. These expressions are thus licensed by virtue of regular principles of syntax. One may therefore expect that the meaning of these PPs is an instance of regular compositional semantics. We will demonstrate that this is indeed the case, adopting the semantic framework of *Lexicalized Flexible Ty2* (LFTy2; (Sailer, 2003)). In this section we will first present LFTy2 and then show how the meaning of CPPs can be computed on the basis of our syntactic assumptions.

2.1 Lexicalized Flexible Ty2

LFTy2 is an adaptation to HPSG of *Flexible Montague Grammar* (Hendriks, 1993). We will take the CONTENT value of a sign to be an expression of a standard semantic representation language, in this case Ty2 (Gallin, 1975). Lexical elements are assigned an expression of Ty2 as their *basic translation*. The CONTENT value of a phrase is the functional application of the CONTENT values of the daughters. In addition, flexible semantic systems provide a number of type shifting operations. These are needed to make the semantic types of sisters compatible with each other, for scope ambiguities and for coordination (see (Hendriks, 1993)). In accordance with (Bouma, 1994) we will apply shifting only to lexical elements. As an illustration, see the PP in (8).

(8) Peter schlief *in einem Hotel*.
Peter slept in a hotel
$\exists x[\textbf{hotel}'(x) \wedge \exists e[\textbf{in}'(x,e) \wedge \textbf{sleep}'(e,\textbf{p})]]$

The semantic derivation of the PP is outlined in Figure 1.4. Every word is assigned a basic translation. The logical form of the NP *einem Hotel* ('a hotel') results from functional application. Since this logical form is of type $(et)t$ it cannot immediately combine with the basic translation of *in* ('in'), $\lambda z_e \lambda R_{e(et)}.\exists e[\textbf{in}'(z,e) \wedge R(u)(e)]$. LFTy2 offers a shifting operation, called AR (*argument raising*), which raises the type of a semantic argument. Here the first semantic argument of *in* is raised to the type $(et)t$ in order to be compatible with the NP.[4]

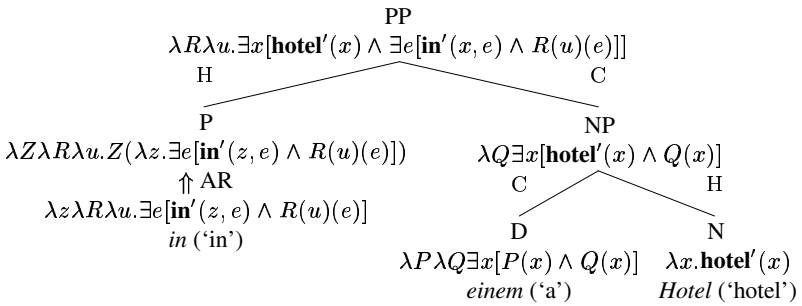

Figure 1.4. The structure of the PP *in einem Hotel* ('in a hotel')

Verbal complexes are the prototypical examples for raising structures, i.e. semantic arguments are not realized as the syntactic complements of the selecting item. Since we plan to analyse CPPs syntactically in analogy with verbal complexes in German, we will first sketch the semantic analysis for verbal complexes. We will then demonstrate that this analysis carries over directly to the PP data. To illustrate this, Figure 1.5 indicates the syntactic structure and the semantic derivation of the VP *Fido füttern will* (*Fido feed want*, 'want to feed Fido').[5]

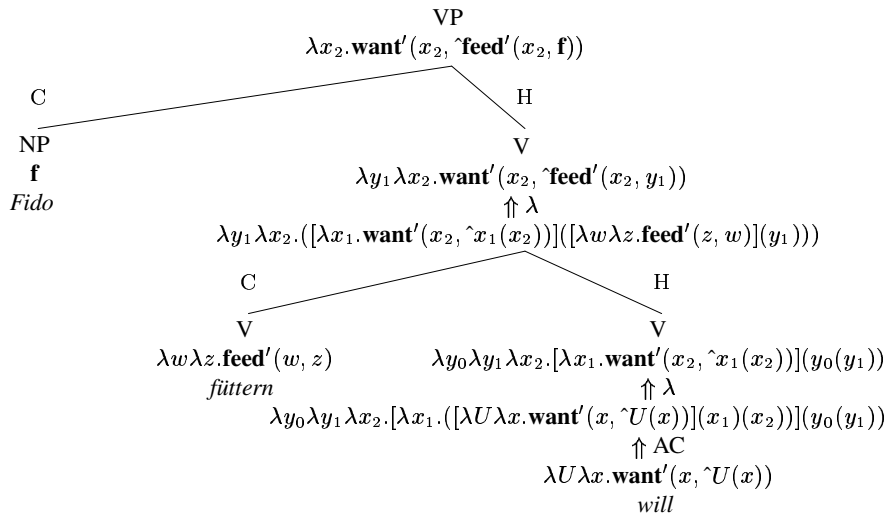

Figure 1.5. The structure of the VP *Fido füttern will* (*Fido feed want*, 'wants to feed Fido')

The LFTy2 fragment in (Sailer, 2003) does not account for syntactic argument raising. In (9) a shifting operation, AC (*argument composition*), is introduced to achieve the correct identification of syntactic constituents and their semantic roles in raising structures. The definition states that if a functor takes an argument of a certain type a_i, it can then combine alternatively with a number of other arguments, which also combine to form an expression of type a_i.

(9) Argument Composition (AC):

AC is a relation between two expressions α and β such that
if α is of some type $a_1(\ldots(a_i \ldots (a_n b)\ldots)\ldots)$, then β is some term $\lambda x_1 \ldots \lambda x_{(i-1)} \lambda y_0 \lambda y_1 \ldots \lambda y_m \lambda x_{(i+1)} \ldots \lambda x_n.[(\lambda x_i.\alpha(x_1)\ldots(x_n))(y_0(y_1)\ldots(y_m))]$
where each x_j is of type a_j, y_0 is of some type $c_1(\ldots(c_m a_i)\ldots)$, and each y_k is of type c_k.

In Figure 1.5, α is the basic translation of *will* ('wants'). For clarity, we have used exactly the same variable names as in the definition of AC. The first semantic argument of *will* ('wants') determines $a_i = et$. y_0 has the type of the bare infinitival complement *füttern* ('feed'), $e(et)$. The direct object of *füttern* ('feed') is syntactically raised and, consequently, its semantic counterpart y_1 appears as an extra argument of type e in the type-shifted expression. This new expression combines with the basic translation of the verb *füttern* ('feed'). As desired, y_0 combines with y_1 to form an expression of type $a_i = et$. For clarity we have indicated the resulting expression before and after λ-conversion ($\Uparrow \lambda$).

2.2 The Meaning of CPPs

We can now address the interpretation of CPPs. We will show that the syntactic structures assumed for these combinations can be interpreted compositionally. To illustrate this we will examine the PP *in Verbindung mit x* ('in connection with x'). We will argue that the words in this combination occur with a logical form which is also available in other combinations, and that the logical form of the overall PP results from the application of shifting operations and functional application as discussed in the previous subsection.

The preposition *in* ('in') occurs in the PP *in Verbindung mit x* ('in connection with *x*') with its metaphorical non-spatial meaning, just as in many other combinations (cf. (10)). For our purpose, we simply assume the same basic translation of *in* ('in') as in Figure 1.4. The preposition *mit* ('with') is used as a selected preposition. Therefore, it does not contribute an independent meaning and is translated as the identity function ($\lambda x.x$). It occurs with this translation in other combinations as well, such as *mit Fisch handeln* (*with fish deal*, 'to deal in fish').

(10) in einer Beziehung/ einer Relation/ diesem Zusammenhang
in a connection/ a relationship/ this context

The noun *Verbindung* ('connection') is a nominalization of the verb *verbinden* ('connect'). The basic translation of the verb is $\lambda z \lambda y \lambda x \lambda e.\mathbf{connect'}(e, x, y, z)$, where e is a "connecting" eventuality, in which x connects y with z. In an HPSG account of *-ung*-nominalizations in German, (Reinhard, 2001) proposes that the suffix *-ung* raises the arguments of the verbal base with which it combines. Which of these arguments can be realized and how they can be realized in syntax depends on the verb class. The example in (11) shows different possibilities of syntactic argument realization.

(11) Eine Verbindung (von bin Laden) mit Hussein wäre absurd.
a connection of bin Laden with Hussein would be absurd
'A connection (of bin Laden) with Hussein would be absurd.'

In (11) the underlying subject of *verbinden* ('connect') remains unexpressed. The underlying direct object is also optional. Unrealized arguments are semantically present but unspecified. Thus we assume that they are existentially bound within the noun's logical form. In (11) this can result in a CONTENT value of the form $\lambda z \lambda e.\exists x \exists y[\mathbf{connect'}(e, x, y, z)]$.

Finally, the PP in (12) has no determiner. The absence of a determiner also has the effect of existential quantification. For further combinatorics, the NP with no determiner must be of type $(et)t$. Thus existential quantification over the referential argument leads to a logical form similar to that of a quantified NP, i.e. to the expression $\lambda z \lambda P.\exists e[\exists x \exists y[\mathbf{connect'}(e, x, y, z)] \wedge P(e)]$ in (12).

(12) Die Raumfähre flog mehrere Tage ohne Verbindung mit der Bodenstation
the space shuttle flew several days without connection with the ground station
durchs All.
through the space
'For several days the space shuttle flew through space without connection to the ground station.'

This is exactly the logical form we need for the interpretation of the CPP *in Verbindung mit x* ('in connection with *x*'). The syntactic structure and the semantic derivation are shown in Figure 1.6. Both are parallel to what is depicted for the VP in Figure 1.5. The basic translation of *in* ('in') first undergoes AR in order to be of the appropriate type to combine with a quantified NP. Then AC is applied and the resulting expression has two semantic argument (y_0 of type $e((et)t)$ and y_1 of type e) instead of the single semantic argument V of type $(et)t$ in the input to AC. This demonstrates that the meaning of the PP can be computed on the basis of independently motivated meaning assignments and shifting rules.

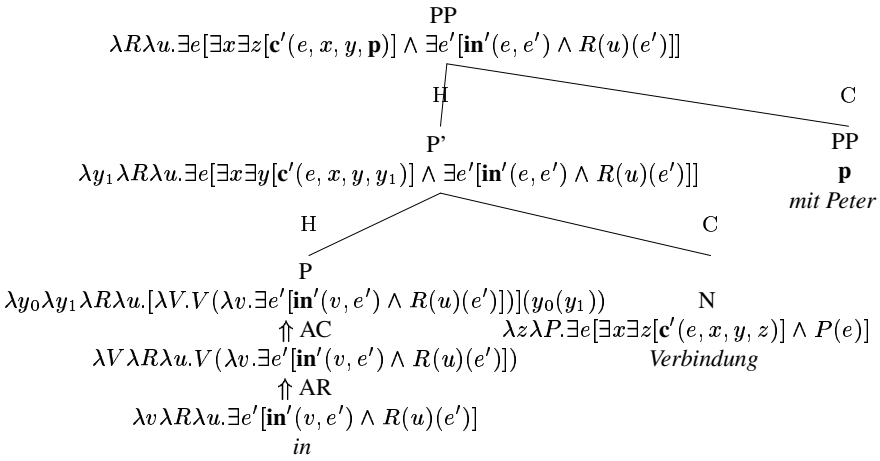

Figure 1.6. The structure of the PP *in Verbindung mit Peter* ('in connection with Peter')

One can treat most $P_1N_1P_2NP$s in which N_1s are deverbal event nominalizations as being semantically regular, i.e. licensed by regular translations and regular derivational operations. This semantic and syntactic regularity explains their high productivity in contemporary German. In the next section we will discuss two types of irregular $P_1N_1P_2NP$ combinations.

3. Irregular Combinations

In this section we will discuss subtypes of CPPs which behave differently with respect to the two *regularity properties* in (13) which are adopted from

(Sailer, 2003).[6] If a given $P_1N_1P_2NP$ sequence lacks at least one of these properties, we will consider it irregular, i.e. of idiomatic character.

(13) **RP1**: Every element of the PP can be attributed a meaning with which it also occurs independent of the combination under consideration.

RP2: The meaning of the entire PP is arrived at by combining the meanings of its parts in a regular way.

If we reconsider the analysis of *in Verbindung mit x* ('in connection with *x*') in Figure 1.6, we see that this CPP shows both regularity properties. Firstly we argued that all the lexical elements in the combination appear with the same meaning assignment in other structures (RP1). Secondly we applied only the rules of syntactic and semantic combination which are independently required in the language (RP2).

Whereas *in Verbindung mit x* ('in connection with *x*') can be described as a fully regular combination, the following two subsections will be devoted to $P_1N_1P_2NP$ combinations which show irregularities with respect to RP1 or RP2. Nevertheless, there are approaches which provide the prerequisites to account for these combinations: external selection (Soehn, 2003) and phrasal lexical entries (Sailer, 2003). We will outline both approaches and show how to apply them to account for the more idiosyncratic CPPs.

3.1 Bound Words

In some irregular $P_1N_1P_2NP$ sequences the N_1 is a so-called *bound word*, e.g. *in Anbetracht von x* ('in consideration of x'). The entire PP is semantically decomposable, and thus satisfies the condition of semantic regularity in RP2. However, RP1 has not been satisfied, since not all components of that PP may occur with the same meaning in other contexts: the noun *Anbetracht* ('consideration') can only occur in combination with the preposition *in* ('in').

To account for bound words within PPs in general, (Soehn, 2003) generalizes the external selection mechanisms of HPSG (cf. the MOD and SPEC features). (Soehn, 2003) assumes that in every type of phrase the non-head daughter can determine syntactic and semantic properties of the head daughter. This idea is realized by conflating the attributes MOD and SPEC into one attribute XSEL (*external selection*), which is appropriate for the sort *head* and takes a *synsem* object as its value. In addition the so-called PRINCIPLE OF EXTERNAL SELECTION (PXS) ensures the identity of the XSEL value of the non-head and the SYNSEM value of the head, similar to the SPEC-PRINCIPLE, which has become obsolete.

In the lexical entry of *Anbetracht* ('consideration') in Figure 1.7, the XSEL value is specified as a *synsem* object with [PFORM *in*]. This specification and the PXS will ensure the occurrence of *Anbetracht* ('consideration') exclusively

within a PP headed by the preposition *in* ('in'). For freely occurring words, the XSEL value is underspecified.[7]

The PP *in Anbetracht von x* ('in consideration of *x*') is a $P_1N_1P_2NP$ expression. Therefore, the XSEL value of *Anbetracht* ('consideration') explicitly requires the preposition *in* ('in') to raise the argument of the bound word, i.e. the PP *von x* ('of *x*') ([2] in the figure).

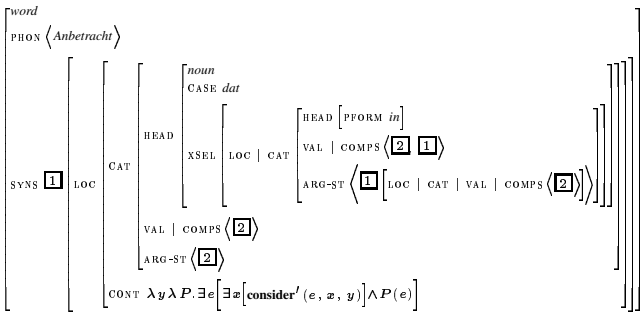

Figure 1.7. The relevant part of the lexical entry of the noun *Anbetracht* ('consideration')

Assuming the usual non-spatial meaning for *in* ('in') selecting *Anbetracht* ('consideration'), we can derive the meaning of the entire PP parallel to the derivation in Figure 1.6. This shows that we can smoothly merge the external selection approach of (Soehn, 2003) with the complement raising approach.

3.2 Phrasal Lexical Items

There are also $P_1N_1P_2NP$ expressions which escape a compositional treatment, such as *an Hand von x* (*at hand of x*, 'by means of x'), *an Stelle von x* (*at place of x*, 'in lieu of x') or *auf Grund von x* (*on base of x*, 'by virtue of x'). This type is significantly less frequent in German than the fully regular combinations. These expressions consist of lexical entities of which each one also appears outside the particular PP. When considering the meaning of any of these PPs it is highly problematic to assign a combination-specific meaning to its particular elements such that the meaning of the entire PP could be derived compositionally. Therefore these combinations do not exhibit RP2. This irregular behavior makes the assumption plausible that these expressions are licensed directly by the lexicon. In this subsection we will provide an analysis of this type of CPPs using the expression *an Hand von* (*at hand of*, 'by means of') as a prototypical example.

In the architecture of (Pollard and Sag, 1994) all syntactically complex signs, i.e. all phrases, are subject to the regular principles of syntactic and semantic combination. However, idiomatic expressions of the type *kick the bucket* ('die') cannot be handled with this kind of appraoch. To overcome this empirical deficiency, (Sailer, 2003) uses *Phrasal Lexical Entries* (PLEs).[8]

(Sailer, 2003) introduces an attribute COLL (*context of lexical licensing*) on the sort *sign*. Signs which are directly licensed by the lexicon have the specification [COLL +], whether they are words or phrases. On the other hand, signs which are licensed by ID schemata or lexical rules have the specification [COLL −]. Consequently there is a LEXICON PRINCIPLE which lists the lexical entries for all signs with a positive COLL value (cf. (14)). This principle contains the usual lexical entries for words (LE_i) as well as phrasal lexical entries for idiosyncratic phrases (PLE_i). Additionally the antecedents of principles of regular combination, such as the ID PRINCIPLE and the SEMANTICS PRINCIPLE, are restricted to phrases with a [COLL −] specification.[9]

(14) The LEXICON PRINCIPLE:

$$\begin{bmatrix} \textit{sign} \\ \text{COLL } + \end{bmatrix} \longrightarrow LE_1 \vee \ldots \vee LE_n \vee PLE_1 \vee \ldots \vee PLE_m$$

We can apply this approach to PPs such as *an Hand von x* (*at hand of x*, 'by means of *x*'). We assume a PLE for the combination *an Hand* (*at hand*, 'by means') which requires a genitive NP or a *von*-PP as its complement. This PLE is outlined in Figure 1.8. It is important to note that even though the phrase *an Hand* (*at hand*, 'by means') is irregular, its daughters *an* ('at') and *Hand* ('hand') occur as exactly the same words in other contexts. However, the semantic contributions of the words are not combined to form the CONTENT of the phrase. Instead, the phrase as a whole receives an idiosyncratic meaning.

$$\begin{bmatrix} \textit{phrase} \\ \text{PHON } \boxed{1} \otimes \boxed{2} \\ \text{SYNS} \begin{bmatrix} \text{LOC} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD } \boxed{3} \\ \text{VAL} \mid \text{COMPS} \left\langle (\text{NP}[\textit{gen}] \vee \text{PP}[\textit{von}]) \right\rangle \end{bmatrix} \\ \text{CONT } \lambda x \lambda R \lambda u. \exists e[\textbf{by-means-of}'(x, e) \wedge R(u)(e)] \end{bmatrix} \end{bmatrix} \\ \text{DTRS} \begin{bmatrix} \text{HEAD-DTR} \begin{bmatrix} \text{PHON } \boxed{1} \langle \textit{an} \rangle \\ \text{SYNS} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD } \boxed{3} \textit{ prep} \end{bmatrix} \\ \text{NHEAD-DTR} \begin{bmatrix} \text{PHON } \boxed{2} \langle \textit{Hand} \rangle \\ \text{SYNS} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD } \textit{noun} \end{bmatrix} \end{bmatrix} \\ \text{COLL } + \end{bmatrix}$$

Figure 1.8. Outline of the phrasal lexical entry of *an Hand* (*at hand*, 'by means')

The use of the phrase *an Hand* (*at hand*, 'by mans') in larger structures is illustrated in Figure 1.9. Note that the COLL values of the phrases *von Prolog* ('of Prolog') and *an Hand von Prolog* (*at hand of Prolog*, 'by means of Prolog') are specified as −, since these phrases are licensed by the regular constraints of grammar. In contrast, the COLL value of the phrase *an Hand* (*at hand*, 'by means') is specified as +. As an internally irregular expression, the phrase *an Hand* (*at hand*, 'by means') is licensed immediately by the lexicon.

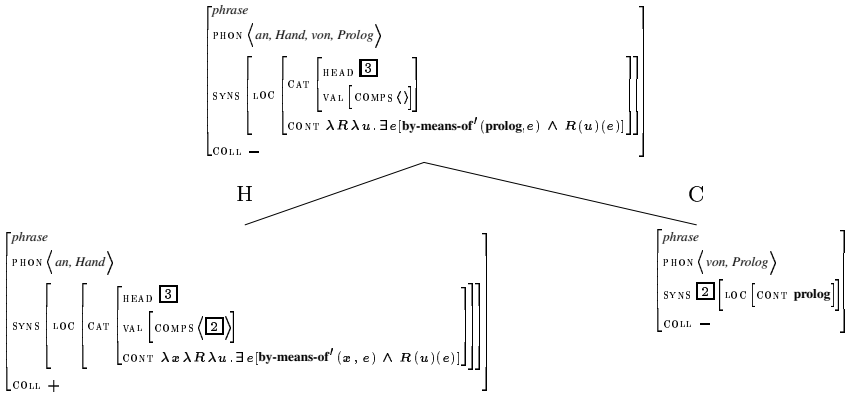

Figure 1.9. The structure of the PP *an Hand von Prolog* (*at hand of Prolog*, 'by means of Prolog')

In this section we have demonstrated that our account of CPPs interacts in an empirically adequate way with HPSG approaches to irregularity phenomena such as the XSEL approach to distributional idiosyncrasies and the phrasal lexical entry approach to combinatorial irregularities.

4. Summary

(Trawiński, 2003) discusses syntactic properties of $P_1N_1P_2NP$ sequences which are the basis for complement raising analysis. Based on this analysis, we have investigated further properties of these CPPs focusing on semantic aspects. The objective of our investigations was to examine whether the semantic representation of these expressions can be derived compositionally. We have thereby seen that the expressions discussed are not uniform with regard to their semantic behavior, forming three classes: CPPs which can be analyzed compositionally (*in Verbindung mit x* ('in connection with *x*')), CPPs involving bound words which can also be treated within the combinatorial semantics but which require some mechanism to describe distributional properties of the particular bound words (*in Anbetracht von x* ('in consideration of *x*')), and CPPs which cannot be handled by virtue of common derivational methods (*an Hand von x* (*at hand of x*, 'by means of *x*')). However, we have shown that the available HPSG approaches, i.e. the external selection approach of (Soehn, 2003) and the phrasal lexical entry approach of (Sailer, 2003), provide the necessary means to account for all of these data.

Acknowledgments

We would like to thank Frank Richter and the anonymous reviewers for their helpful comments. We are also grateful to Guthrun Love for her help with the proofreading in English.

Notes

1. Cf. (Lindqvist, 1994), (Quirk and Mulholland, 1964), (Beneš, 1974), etc.

2. It is unclear how many $P_1\,N_1\,P_2$ expressions there are in German. (Schröder, 1986) identifies more than 90. (Beneš, 1974) 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 CPPs in German see also (Meibauer, 1995).

3. For further applications of the raising mechanism whithin the HPSG grammar framework see e.g. (Hinrichs and Nakazawa, 1989), (Hinrichs and Nakazawa, 1994), (Meurers, 2000) or (De Kuthy, 2000).

4. We deliberately simplify the treatment of the eventuality variable *e* when we assume that the quantifier which binds *e* is introduced by the preposition. This simplification has no bearing on the main argumentation in this paper.

5. We have left out the eventuality variables in this example for simplification.

6. (Sailer, 2003) applies analogous criteria to VPs. There, expressions with *bound words* such as *make headway*, violate the first regularity property. Non-decomposable idiomatic expressions such as *kick the bucket* show the corresponding violation of the second property.

7. Nouns often show idiosyncratic preferences for a particular preposition, such as *in / *bei Verbindung mit ('in / *at connection with')*. The XSEL approach can capture this. For the lexical entry of *Verbindung* ('connection') we only have to add a constraint stating that if the XSEL value of the noun is a raising preposition, then this preposition has the PFORM value *in* ('in'). The same solution can be applied to the noun-specific choice of support verbs.

8. For an alternative constructional approach to idioms see (Riehemann, 2001) or (Sag et al., 2002).

9. For phenomena discussed in this paper it is sufficient to adopt the simplified usage of the COLL attribute as presented in Section 8.1 of (Sailer, 2003). In his Section 8.3 (Sailer, 2003) assumes that the COLL attribute takes a list of signs as its value, such that for every non-lexical sign, the COLL value is an empty list. The COLL value of a lexical sign is a singleton list containing the root sign of the utterance in which this lexical sign occurs. With this more complex mechanism, CPPs with bound words can also be described. However, as elaborated in (Soehn and Sailer, 2003), this more general use of COLL might be too powerful.

References

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

Bouma, G. (1994). Calculated Flexibility. In Bunt, H., Muskens, R., and Rentier, G., editors, *Proceedings of the International Workshop on Computational Semantics*, pages 32–40. Katholieke Universiteit Brabant.

De Kuthy, K. (2002). *Discontinuous NPs in German — A Case Study of the Interaction of Syntax, Semantics and Pragmatics*. CSLI Publications, Stanford.

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

Gallin, D. (1975). *Intensional and Higher-Order Modal Logic*. North-Holland, Amsterdam.

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

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

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

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

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

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

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

Quirk, R. and Mulholland, J. (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*, pages 64–73, Amsterdam.

Reinhard, S. (2001). *Deverbale Komposita an der Morphologie-Syntax-Semantik-Schnittstelle: ein HPSG-Ansatz*. PhD thesis, Universität Tübingen.

Riehemann, S. Z. (2001). *A Constructional Approach to Idioms and Word Formation*. PhD thesis, Stanford University.

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

Sailer, M. (2003). *Combinatorial Semantics and Idiomatic Expressions in Head-Driven Phrase Structure Grammar*. PhD thesis (2000), Universität Tübingen. Published as: Arbeitspapiere des SFB 340, Nr. 161.

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

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 Sailer, M. (2003). At First Blush on Tenterhooks. About Selectional Restrictions Imposed by Nonheads. In Jäger, G., Monachesi, P., Penn, G., and Wintner, S., editors, *Proceedings of FGVienna: The 8th Conference on Formal Grammar*, pages 149–161. To appear also as CSLI Publications Online Proceedings.

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

Trawiński, B. (2003). A New Application for Raising in HPSG: Complex Prepositions. In Jäger, G., Monachesi, P., Penn, G., and Wintner, S., editors, *Proceedings of FGVienna: The 8th Conference on Formal Grammar*, pages 163–175. To appear also as CSLI Publications Online Proceedings.