The annotation of preposition senses in German

Antje Müller, Claudia Roch, Tobias Stadtfeld, Tibor Kiss

1. Treatment of prepositions and their polysemy

1.1. Missing findings on preposition senses?

It seems to be common wisdom that prepositions (especially simple prepositions) are highly polysemous. Preposition senses seem to be well explored, as there exists plenty of literature on prepositions and their interpretations. On a closer inspection, however, it turns out that we are far away from a well-structured understanding of preposition senses.

There is mutual consent about simple preposition being derived from adverbs and retaining the basic meanings of those adverbs, like a temporal, spatial or modal usage. Thus, most simple prepositions in German have spatial interpretations, from which metaphorical usages have been derived. Temporal interpretations are also common with the majority of simple prepositions, whereas other interpretations (as e.g. existence – cf. chapter 5) are represented to a lesser extent. Precisely those senses are the ones being poorly understood and studied to a lesser extent. But even for the well examined ones, there is the need for a representational pattern suitable not only for a few but for all prepositions at best. This demand gains particular importance if one tries to determine the distribution of preposition senses in large corpora. An annotation scheme is required in the first place to automatically classify senses or to determine the distribution of contextual patterns of preposition senses in annotated corpora.

1.2. Existing literature on preposition senses

Existing works on preposition senses can be classified in more or less four classes. The first one contains dictionaries, either for native speakers or for foreign-language learners. Those dictionaries do not only cover prepositions but “all” words of a given language. By implication, the information on prepositions can't be very detailed. Quite often, the description of prepositions is limited to some simple – often artificially designed – examples
and some more or less informative labels for the senses. Criteria for the
distinction of senses are not provided and the interpretation of sense labels
varies across the prepositions described.

The second class contains dictionaries or collections of prepositions in
particular, often with the attempt to systematize the spectrum of preposition
meanings. Schröder's “Lexikon deutscher Präpositionen” (Schröder 1986)
stands out in including a fine-grained feature-based analysis of preposition
senses in German. Making use of over 200 binary features, it is, however,
too complex for the demand of corpus annotation.

Reviewing the existing grammars and dictionaries, it is conspicuous that
the examples used in the reference works are repeated over and over again,
rarely being revised or replaced. As the characterization of possible senses
is often carried out in neglect of large data sets, senses tend to be ignored
that do not come to mind immediately.

The third class of existing works contains studies to a single preposition,
some in one language, some across languages. The advantage here lies in a
possibly much more detailed analysis, but existing connections to other
prepositions and their senses cannot be captured. Consider e.g. the condi-
tions for interchanging prepositions with identical/similar senses, as illus-
trated with nach and gemäß in (1).

\[(1)\] Tiere sind nach deutschem Recht eine Sache.
Tiere sind gemäß deutschem Recht eine Sache.
Animals are according to German law a thing.
‘According to German law animals are things.’

The fourth class picks out a sub group of prepositions with a common
sub sense. There are publications concerning for example spatial or tem-
poral prepositions (neglecting that it is only a temporal or spatial use of the
preposition), describing the options of realization of space and time in lan-
guage. They seldom touch on the other sub senses of the spatial or temporal
prepositions.

2. Preposition senses and preposition-noun combinations

Our interest in dealing with preposition senses originates with our research
on preposition-noun combinations (PNCs). Hence, we present an annota-
tion schema for preposition senses for prepositions occurring in PNCs and
PPs in German. PNCs are combinations of prepositions with determinerless nominal projections such as in (2):

(2) *auf Anfrage* (‘after being asked’), *unter Androhung* (‘under threat’),
    *mit Vorbehalt* (‘with reservation’)

PNCs can be extended, as long as they do not contain a determiner. Pre-nominal modification is permitted (3), as well as postnominal complementation (4).

(3) *auf parlamentarische Anfrage* (‘after being asked in parliament’), *mit beladenem Rucksack* (‘with loaded backpack’)

(4) *Er wehrt sich gegen die Forderung nach Stilllegung einer Verbrennungsanlage.*
    ‘He defies the demand for closing an incineration plant.’

PNCs present an anomaly in the grammar: they violate the rule that countable singular nouns have to appear with a determiner. For some time, PNCs have been treated as exceptions, but recent research has shown that they are indeed productive and no more idiomatic than other phrasal combinations (Stvan 1998; Baldwin et al. 2006 for English; Dömges et al. 2007; Kiss 2007 for German).

It has been claimed that either the semantics of the noun or the semantics of the preposition (or both) might play a major role in determining whether a singular count noun may appear without a determiner in a PNC.

Baldwin et al. (2006) for example have claimed that preposition senses in PNCs in English are more restricted than in PPs in general. Initial investigations have shown that senses of prepositions in German PNCs are restricted, too. The preposition *unter* (‘under’), for instance, does not allow spatial interpretations in PNCs, unless it appears in the context of newspaper headlines, but there is no evidence for a general ban on local interpretations in PNCs.

The conditions for determiner omission, however, have not been detected yet, and conditions applying to one language can't be assumed to carry over to other languages. In addition, speakers only reluctantly judge the acceptability of newly coined PNCs, so that reliance to introspective judgments cannot be assumed.

We propose an analysis for PNCs that is based on corpus annotation, annotation mining (Chiarcos et al. 2008), and logistic regression modeling.
Antje Müller, Claudia Roch, Tobias Stadtfeld, Tibor Kiss

(Harrell 2001). Hence, linguistically relevant generalizations can be derived in a bottom-up fashion from a suitably annotated corpus with the help of annotation mining. Annotation mining describes the process of adding multiple layers of annotation to an existing corpus and to derive generalizations from the layers of annotation. Annotation mining makes use of part-of-speech tags, shallow and deep syntactic analysis, sense annotation, and complex conceptual annotations derived from resources such as HaGenLex (Hartrumpf, Helbig, and Osswald 2003) and GermaNet (Kunze and Lemnitzer 2002). Taken together, the annotations serve to define licensing conditions for the omission of a determiner in PNCs.

But not every preposition can appear in a PNC. Hence, we limit our analysis to the (simple) prepositions that are allowed in PNCs and typically take NP complements when appearing in PPs – the number in parentheses offers the number of possible senses for each preposition:

(5) mit (12), unter (11), für (9), über (9), in (8), auf (7), an (6), bei (6), gegen (6), nach (6), ohne (6), um (6), durch (5), vor (5), neben (3), dank (2), hinter (2), mittels (2), wegen (2), gemäß (1), seit (1), während (1)

The annotation of preposition senses forms part of the larger task of annotating a corpus on all possibly relevant levels. But it is clearly one of the most challenging parts, not at least because an annotation schema for preposition senses in German did not exist. Such an annotation schema should obviously not only be useful for the task at hand, but should also more generally allow the definition of a reference corpus for automated preposition sense annotation. Hence the schema has to be feasible for manual annotation, to begin with.

3. Building an annotation schema for preposition senses

3.1. The reference corpus

For our annotation work, we use a newspaper corpus of the Swiss-German newspaper “Neue Zürcher Zeitung” from 1993 to 1999, comprising approx. 230 million words. The annotation is based on an XML-stand-off format. The annotation tool MMAX2 (Müller and Strube 2006) is used for manual annotation.
For each preposition, we consider three datasets: PNCs, where N is a count noun, corresponding PPs with the same count noun, and PPs containing count nouns not appearing inside PNCs. The identification of count nouns is based on the combination of two classifiers. The analysis of the classifiers resulted in 4431 fully countable nouns that form the basis for PPs and PNCs in the three datasets. For each dataset in the corpus the following annotations are provided:

- **Lexical level**: Part-of-speech, inflectional morphology, derivational morphology of nouns, interpretation of nouns, interpretation of prepositions, noun compounding.
- **Syntactic level**: Mode of embedding of the phrase (adjunct or complement), syntactic dependents of the noun, modification of the noun.
- **Global level**: Is the phrase contained in a headline, title, or quotation? Is the phrase idiomatic?

Headlines, titles, and quotations are particularly prone to text truncation. Similarly, idiomatic PNCs and PPs might follow combination rules that differ from the general modes of combination.

For automatic annotation the following tools are applied: The Regression Forest Tagger (Schmid and Laws 2008) for POS tagging and morphological analysis (the tagger contains the SMOR component for morphological analysis, cf. Schmid 2004), the Tree Tagger (Schmid 1995) for chunk parsing, and the Malt-Parser (Nivre 2006) for syntactic dependencies.

To determine noun meanings we make use of two resources. The first resource is GermaNet (Kunze and Lemnitzer 2002), the German version of WordNet. We employ 23 top-level categories, and each noun is annotated with every top-level category it belongs to. Secondly, we use the computer lexicon HaGenLex (Hartrumpf, Helbig, and Osswald 2003), which offers specific sortal information derived from a formal ontology for each noun.

As the current set-up is concerned with the creation of a reference corpus, the annotations on the global level, as well as the preposition sense annotation are carried out manually.

### 3.2. An inventory of preposition senses

Currently, a standardized inventory of preposition senses does not exist for German. Ideally, one would employ a universally applicable schema, but such a schema would require language-specific mappings from senses to prepositions and vice versa. The Preposition Project (cf. Litkowski and Hargraves 2005) offers categories for preposition senses in English, and
thus could serve as a starting point for the development of a similar schema in German. However, we considered it easier to develop a schema from scratch than to transfer English sense-preposition pairs to German prepositions and their senses.

We started by consulting the German grammar by Helbig and Buscha (2001) and the dictionary “Duden Deutsch als Fremdsprache” (Duden 2002), as well as Schröder’s aforementioned dictionary of German prepositions. For spatial and temporal senses additional resources were consulted because the information provided by the dictionaries was insufficient.

Dictionaries quite often do not strive for complete coverage of preposition senses. This may originate from insufficient consideration of corpus data. So we did not only rely on the reference works and determined our schema with the help of examples collected there, but defined the schema iteratively, testing it against corpus data. We used a temporary state of the schema for the annotation of corpus data to detect missing senses, discrepancies between interpretations or grubby classifications. If necessary the schema was changed and the annotated data reworked.

The schema makes use of three characteristic features. First, the annotation schema is hierarchically organized, allowing for the inclusion of taxonomies of sub senses and the assignment of intermediate senses to instances. Secondly, the schema does not prohibit the assignment of more than one sense to an instance. This property is useful if further disambiguation proves to be infeasible. Thirdly, certain general properties of senses are extracted as cross-classifying features, allowing a cleaner representation of the sense taxonomy. A prototypical application of the last feature concerns the ubiquitous distinction between local and directional senses.

As was already mentioned, the schema is formally based on a hierarchical tree-like structure, catering for sub hierarchies of individual senses. Beginning with a root node, types of preposition meanings branch to sub trees for different classes (e.g. local, temporal or causal) with differing depths or to individual, non-splitting branches (cf. Figure 1).

From our analysis of grammars and dictionaries, we have identified 27 different top level interpretations for our restricted inventory of 22 prepositions. It should not come as a surprise that the senses SPATIAL, TEMPORAL, CAUSAL, and MODAL are mapped to sub hierarchies of senses. In addition, we assume that the sense EXISTENCE – which is only employed by the prepositions ohne and mit (cf. section 5) – also takes two subcategories, viz. ANALYTIC and SYNTHETIC. The schema is generally agnostic to the ontological status of (intermediate) super senses. Assignment of a super sense
instead of a most specific sub sense may thus imply that a more specific interpretation is conceivable but cannot be derived from the criteria at hand.

While the schema includes an automatic mapping from sub senses to super senses, it is naturally desirable that the most specific annotation is provided whenever possible. For spatial and temporal senses, the determination of most specific sub senses is guided by decision trees implemented in MMAX2 (cf. next section).

The remaining senses do not show sub senses and are typically only instantiated by few prepositions, like for example AGENT or EXTENSION. A full list of the top-level interpretations together with the number of prepositions that can instantiate them is given below.

\[(6)\] causal (15), local (14), temporal (14), modal (13), center of reference (8), affiliation (5), order (5), exchange (4), comparison (3), statement (3), theme (3), adversative (2), communality (2), inclusive (2), correlation (2), distributive (2), existence (2), hierarchy (2), participation (2), recipient (2), state (2), substitute (2), transgression (2), agent (1), copulative (1), extension (1), realization (1).

In some cases, a decisive distinction between two senses cannot be drawn:

\[(7)\] Fire after/because of lightning stroke

It is for cases like (7) that the schema allows the assignment of multiple senses to individual instances.

In addition to semantic features, the schema contains the feature GOVERNED for prepositions governed by a lexical head. Governed prepositions are often considered to show light semantics only, if at all. But the assign-
ment of the feature GOVERNED does not preclude the assignment of additional semantic features if it turns out that the preposition shows a discernible meaning despite its being governed.

On a quantitative level, three types of relations between prepositions and senses can be distinguished: The first group comprises of highly polysemous prepositions (such as mit, unter, für, über, in, auf) with up to 11 discernable top-level senses. It is a characteristic trait of the second group – which comprises many of the 22 prepositions under investigation – that they exhibit polysemy w.r.t. local, causal, temporal, and modal senses. Examples are auf or nach, which both instantiate local, causal, temporal and modal senses. Finally, many other senses are only instantiated by few prepositions, like SUBSTITUTE, AGENT or EXTENSION. It would thus be premature to call prepositions highly polysemous in general.

3.3. Properties of spatial and temporal decision trees

For sense labels like TEMPORAL or LOCAL, it seems obvious that just assigning the sense is not particularly revealing. Yet, examples in dictionaries often show this kind of characterization, leaving open questions e.g. as to why two prepositions with local senses and identical case requirements cannot be substituted salva veritate.

Instead, we assume structured sub senses of senses like SPATIAL and TEMPORAL, where the assignment of specific sub senses is guided by decision trees, which also facilitate the annotation process.

3.3.1. Spatial interpretations

The spatial decision tree entails the original spatial interpretations of the prepositions as well as some metaphorical extensions, which occur on a regular basis. With regard to metaphorical extensions of spatial senses, our analysis relies on the description in Schröder (1986); for non-metaphorical senses, however, modifications were required as Schröder makes use of several distinctions that deem us unnecessary or counter-intuitive. As an illustration consider Schröder’s distinction between a basal and a surface plane as reference points. It is not possible to determine the interpretation of spatial prepositions like über (‘above’) or unter (‘under’) in some contexts given a body with both planes, as in (8).

(8) Ein Stromkabel verläuft über dem Haus.
    ‘A power supply line runs above the house.’
We do not want to commit ourselves, whether the cable runs above a surface plane, a basal plane, or even above both planes in examples like (8). Instead of defining the (part of the) reference object concerned, we want to define the relation between the reference object and the object to be localized. This is the relation the preposition reflects in the original spatial uses. Therefore we make use of three given reference axes, two horizontal (sidewise and behind/ahead) and a vertical one given by gravity (above/below). Additionally, we use the opposite pair inside/outside, as shown in Figure 2.

Figure 2. Reference axes

It might be apparent to the reader that no distinction between local and directional readings has yet been made. Instead of adopting a top-level distinction, we are able to capture the difference systematically within our spatial decision tree. Our perspective corresponds to theories on the composition of local and directional terms, like Jackendoff (1983). We make use of a cross-classifying feature DIRECTIONAL, applicable to all spatial senses and thus being similar to Jackendoff’s PATH-function (Jackendoff 1983:163).

Specifically, DIRECTIONAL is one of three cross-classifying features. It is applicable to all leaves of the spatial decision tree in Figure 3. With the exception of a single preposition (nach), the investigated spatial interpretations show a case alternation corresponding to the difference between local and directional interpretations, so that the decision between local and directional can be based on the case of the complement.

The other two cross-classifying features used are CONTACT and TANGIBLE. CONTACT is a feature applying to all leaves in the reference domain, signifying whether or not a contact has been established. The feature TANGIBLE specifies the concreteness of the domain. As an illustration, consider Himmel (‘sky’) in (9) which surely is ABOVE but is not tangible.

As (10) shows, all three features can appear together.
(9)  Sterne am Himmel (‘stars in the sky’)

(10)  Er fährt gegen die Wand. (‘He drives against the wall.’)

Figure 3. Spatial decision tree
3.3.2. Temporal interpretations

With regard to the temporal interpretations, we have been able to build on a decision tree for temporal interpretations of German prepositions developed in Durell and Brée (1993). The decisions in the tree are based on the distinction between a matrix and a subordinate eventuality, the characteristics of these eventualities as well as on the identification of the temporal relationship between them.

Since our concern was to obtain a useful feature space to annotate the temporal interpretations of prepositions as they are used in PNCs or PPs (and hence we have a different goal than Durell and Brée's) there was no need for the features included to be distinct for all prepositions. That gave us the room to modify the tree of Durell and Brée on some points. On other points we had to amend Durell and Brée's tree because it did not consider certain interpretations.

The tree itself is subdivided in two parts: One for prepositions building a time measure phrase with their complements, and one for the ones who don't. A time measure phrase can either frame the duration of an eventuality (the matrix eventuality) (11) or define the time of taking place of the matrix eventuality in connection to a reference time (12) (which is equal to the utterance time in this example).

(11) *Sie wohnt seit drei Jahren in Hamburg.*

She lives for three years in Hamburg

‘She is living in Hamburg for three years now.’

(12) *Vor einer Woche haben die Kurse angefangen.*

Ago one week have the courses started

‘The courses started one week ago.’

Non-time-measure readings are found in the upper half of the tree in Figure 4. They relate two eventualities to each other, the matrix eventuality and the sub eventuality. It is defined whether they occur at the same time or in sequence, and whether they label points in time or periods.

In example (13) the matrix eventuality is *gehen wir schlafen* (‘we went to sleep’) and the sub eventuality is *nach dem Essen* (‘after dinner’). They do not occur at the same time but the subevent took place in time before the matrix event.

(13) *Nach dem Essen gehen wir schlafen.* (‘After dinner we went to sleep.’)
Figure 4. Temporal decision tree
4. Inter-annotator-agreement

To determine the feasibility of an annotation schema, usually a weighted kappa statistic ($\kappa$) is employed. But there are two aspects of the present annotation schema that prohibit a direct application of a statistic suggested by Artstein and Poesio (2008), based on Cohen's seminal work.

First, the annotation schema makes use of a hierarchy with subtypes, which leads to overlapping annotation categories. As an illustration, assume that one annotator has annotated a given preposition with the sense MODAL, while a second annotator makes use of the annotation MEDIAL, the latter being a subtype of the first. Secondly, the annotation schema allows more than one annotation for the same token, to cover cases where an ambiguous interpretation cannot be maximally reduced.

To deal with the first problem, Müller et al. (2010) propose to include the hierarchical structure of the annotation schema in the calculation of the weight coefficients for $\kappa$. Basically, two annotations are more closely related if either both annotations are dominated by the same set of nodes in the hierarchy, or one annotation is a direct subtype of the other one. Accordingly, the weight coefficient for a given disagreement is reduced in relation to the depth of embedding of the subcategories, based on the cardinality of the set of nodes that dominate both categories.

With regard to multiple annotations, the weight of disagreement is determined as the arithmetic mean of the cross product of disagreeing annotations. For a detailed explanation, some illustrating examples and the formulas see Müller et al. (2010).

Now, instead of determining the $\kappa$ statistic on the basis of non-overlapping, i.e. mutually exclusive categories, the weights are determined by taking the tree structure into account. Based on the weighted kappa statistic, we have carried out an evaluation based on 1,336 annotated tokens, using the prepositions *an, auf, bei, neben, unter*, and *vor*. The following table summarizes the results for the full set of sense annotations, as well as for senses with subtypes (LOCAL, TEMPORAL, CAUSAL, MODAL), as well as for some individual senses.
Table 1. Subset of weighted kappa-values

<table>
<thead>
<tr>
<th>sub tree with the following root node</th>
<th>$K_w$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROOT</td>
<td>0.657</td>
</tr>
<tr>
<td>local</td>
<td>0.723</td>
</tr>
<tr>
<td>causal</td>
<td>0.575</td>
</tr>
<tr>
<td>modal</td>
<td>0.551</td>
</tr>
<tr>
<td>temporal</td>
<td>0.860</td>
</tr>
<tr>
<td>local_reference_plane</td>
<td>0.603</td>
</tr>
<tr>
<td>temporal_M=S</td>
<td>0.854</td>
</tr>
</tbody>
</table>

The overall result of 0.657 provides support for the general feasibility of the annotation schema, and the results for local and temporal senses are particularly promising. The results for modal and causal senses indicate the necessity to take a look at the data again and to identify sources of error.

5. A closer look – the preposition ohne (‘without’)

A closer look at the preposition ohne (‘without’) will serve to illustrate the annotation schema. This preposition is particularly interesting not only in the context of PNCs but also in determining preposition senses in general. The preposition ohne is the only preposition that appears more often in PNCs than in PPs – and this is an observation that carries over to other languages as well, as has been reported for its Dutch, French, and English counterparts in de Swart et al. (2010). The preposition shows six discernible top level senses (with 13 total sub senses), and thus appears to be quite polysemous, but three of the six top level senses belong to the core interpretation of ohne, and are not found with other prepositions apart from ohne’s cognate mit (‘with’). These three senses are (lack of) EXISTENCE, (lack of) PARTICIPATION, and (not) COMITATIVE. It should be noted that ohne in general expresses a privative meaning, i.e. it expresses the absence or lack of something that is given or present in the analog interpretation of mit. Hence, all senses of ohne are to be interpreted as negations of the usual senses. In addition to the three aforementioned interpretations, the following sub senses of ohne can be identified: MODAL (MANNER, CONCURRENT CIRCUMSTANCE, MEDIAL, INSTRUMENTAL), CAUSAL (IN THE NARROWER SENSE, CONDITIONAL), and (not) INCLUSIVE.

Modal interpretations describe the mode of an event in the broader sense. This might be the way of doing something or a concomitant circum-
stance of an event. Instrumental interpretations also fall in the modal category – using an instrument is just one manner of doing something.

The subcategory MODAL – MANNER describes the manner of an event (14). It includes intention and is influenceable by an agent. In contrast to this, CONCURRENT CIRCUMSTANCE describes external circumstances or a mode of an event, that is not controlled by an agent and there is no intention included (15).

(14) Eine Mofalenkerin, die ohne Helm unterwegs war,…

‘A motorcyclist, riding without a helmet…’

(15) Das nachgebaute Skelett steht aufrecht und ohne Makel in der Landschaft.

‘The copy of the skeleton is standing in the landscape upright and without any defect.’

Medial interpretations describe an aid for an event but no instrument in the narrower sense (16). Instruments in the narrower sense are instruments intentionally used by an agent. They must be concrete objects with a clear instrumental character (17). Medial interpretations, however, cover instrumentalized events as well as all kinds of means to an end. Often, a replaceability with via, by means of or with the aid of is given.

(16) Der Text beschreibt einen Geist der sich – ohne das äussere Korrektiv allgemeiner Werte – die Welt neu erklärt.

‘The text describes a ghost, who newly defines the world for himself without the external corrective influence of universal values.’
Der italienische Sänger demonstriert, dass seine sonore Stimme auch ohne Mikrophon wirkt.

‘The Italian singer demonstrates that his sonorous voice is also effective without a microphone.’

The interpretation PRESENCE describes the lack of an object or person or of an attribute or characteristic of an object or person (18). Its two sub interpretations are ANALYTIC and SYNTHETIC. Analytic interpretations describe mereological relations, for ohne a part of a larger structure is missing (19). The interpretation SYNTHETIC in contrast describes the lack of a defining but not essential attribute. This can be paraphrased with exclusive.

Teenager without an apprenticeship position

‘Teenager without an apprenticeship position’

‘No See TV’ – das Fernsehen ohne Bild

‘No See TV’ – the television without picture

‘No See TV’ – television without a video signal

Mass without a sermon

‘Mass without a sermon’

The interpretation PARTICIPATION describes the lack of an active or passive participation in an action or event.

Hearings without the accused

‘Hearings without the accused’

The causal interpretations express relations found in a chain of causation. This might be relations of reason, cause, purpose or concession. From the causal interpretations ohne instantiates CAUSAL IN THE NARROWER SENSE and CONDITIONAL.

The causal interpretation in the narrower sense gives a cause or reason for something. It can mostly be paraphrased with because of.
The annotation of preposition senses in German

(22) Ohne die absolute Parlamentsmehrheit hat sie noch kein klares Mandat.
Without the absolute parliamentary majority has she yet no clear mandate
‘Without the absolute parliamentary majority she does not have a clear mandate yet.’

Conditional interpretations can be paraphrased with if…then but one must consider the negation in the case of ohne.

(23) Ohne Basiswissen bleibt die Innovation auf der Strecke.
Without basic knowledge stays the innovation on the route
‘Without basic knowledge, innovation falls by the wayside.’

The interpretation INCLUSIVE is restricted to occurrences with numeral adjectives and gives an exclusion of something or someone in the case of ohne.

(24) Ohne den Lehrer waren sie zu zehnt.
Without the teacher were they in tens
‘Without/Excluding the teacher there were ten of them.’

The comitative interpretation of ohne connotes ‘without the company of’. It is an exclusion of someone from an action or event.

(25) Sie ging ohne Hund in die Stadt.
She went without dog in the city
‘She went to town without her dog.’

6. Summary

Although there exists much literature on preposition senses, no annotation schema for the interpretation of German prepositions is available. We have developed such a schema for the manual annotation of preposition senses in German PNCs and PPs. The annotation will lead to a reference corpus of preposition senses in PPs and PNCs. The corpus will not only be useful as a resource for further investigating the realization of preposition senses in PPs and PNCs, but can also be used as a reference corpus for training automatic methods for preposition sense tagging.

An illustration for the interpretations used and their distinction is given using the example of the preposition ohne.
7. Notes

1. Considering the 22 simple prepositions we investigate here, it turns out that 14 of them show spatial interpretations.

2. Nouns that are assigned to more than one top-level category are presumably homonymous or polysemous. We do not disambiguate the nouns. The reason is that individual features will be evaluated for their effect in a logistic model, and an ambiguous noun will receive a value in each feature. Hence, we can be sure that a significant semantic feature will be included in the classification.

3. To distinguish the two sub interpretations it can helpful to replace ohne with mit. If there is no further modification of an analytic feature the resulting construction with mit sounds somehow odd because the specification given is redundant.

   1) Hand ohne Daumen
   2) ?Hand mit Daumen.
   3) Hand mit großem Daumen.

References

Artstein, Ron, and Massimo Poesio

Baldwin, Timothy, John Beavers, Leonoor van der Beek, Francis Bond, Dan Flickinger, and Ivan A. Sag

Chiarcos, Christian, Stefanie Dipper, Michael Götze, Ulf Leser, Anke Lüdeling, Julia Ritz, and Manfred Stede

de Swart, Henriette, Bert Le Bruyn, and Joost Zwarts.

Dönges, Florian, Tibor Kiss, Antje Müller, and Claudia Roch
The annotation of preposition senses in German

Duden

Durell, Martin, and David Brée

Harell, Frank E.

Hartmann, Sven, Hermann Helbig, and Rainer Osswald

Helbig, Gerhard, and Joachim Buscha

Jackendoff, Ray

Kiss, Tibor

Kunze, Claudia, and Lothar Lemnitzer

Litkowski, Ken, and Orin Hargraves

Müller, Antje, Olaf Hülscher, Claudia Roch, Katja Keßelmeier, Tobias Stadtfeld, Jan Strunk, and Tibor Kiss

Müller, Christoph, and Michael Strube

Nivre, Joakim
Schmid, Helmut

Schmid, Helmut, Arne Fitschen and Ulrich Heid

Schmid, Helmut, and Florian Laws

Schröder, Jochen

Stvan, Laurel Smith
1998  The Semantics and Pragmatics of Bare Singular Noun Phrases. Ph. D. diss., Northwestern University, Evanston/Chicago, IL.