An Annotation Schema for Preposition Senses in German

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

Sprachwissenschaftliches Institut, Ruhr-Universität Bochum
D-44801 Bochum, Germany
{mueller, roch, kesselmeier, stadtfeld, strunk, tibor}
@linguistics.rub.de, olaf.huelscher@rub.de

Abstract

Prepositions are highly polysemous. Yet, little effort has been spent to develop language-specific annotation schemata for preposition senses to systematically represent and analyze the polysemy of prepositions in large corpora. In this paper, we present an annotation schema for preposition senses in German. The annotation schema includes a hierarchical taxonomy and also allows multiple annotations for individual tokens. It is based on an analysis of usage-based dictionaries and grammars and has been evaluated in an inter-annotator-agreement study.

1 Annotation Schemata for Preposition Senses: A Problem to be Tackled

It is common linguistic wisdom that prepositions are highly polysemous. It is thus somewhat surprising that little attention has been paid to the development of specialized annotation schemata for preposition senses.[1] In the present paper, we present a tagset for the annotation of German prepositions. The need for an annotation schema emerged in an analysis of so-called Preposition-Noun Combinations (PNCs), sometimes called determinerless PPs or bare PPs. PNCs minimally consist of a preposition and a count noun in the singular that appear without a determiner. In (1), examples are given from German.

(1) *auf parlamentarische Anfrage* (after being asked in parliament), *bei absolut klarer Zielsetzung* (given a clearly present aim), *unter sanfter Androhung* (under gentle threat)

The preposition-sense annotation forms part of a larger annotation task of the corpus, where all relevant properties of PPs and PNCs receive either automated or manual annotations. In developing an annotation schema for preposition senses, we pursue two general goals:

I. An annotation schema for preposition senses should provide a basis for manual annotation of a corpus to determine whether the interpretation of prepositions is a grammatical factor.
II. The preposition sense annotations together with the other annotations of the corpus should serve as a reference for the automatic classification of preposition senses.

With regard to the goals formulated, the present paper is an intermediate report. The annotation schema has been developed and the manual annotation of the corpus is well under way. The next logical steps will be to apply the annotations to a wider range of prepositions and eventually to use the annotated corpus for an automated classification system for preposition senses.

As PNCs form the basic rationale for the current investigation, we are only considering prepositions that occur in PPs *and* PNCs in German. We thus systematically exclude prepositions that do not take an NP complement, postpositions, and complex prepositions. Thus, the sense annotation for prepositions currently comprises the following 22 simple prepositions in German:

(2) an, auf, bei, dank, durch, für, gegen, gemäß, hinter, in, mit, mittels, nach, neben, ohne, seit, über, um, unter, vor, während, wegen

As empirical base of the analysis, we use a Swiss German newspaper corpus, which contains about 230 million tokens (Neue Zürcher Zeitung 1993-1999).

The remaining paper is structured as follows: Section 2 is devoted to the characteristics of the annotation schema. In section 3, we present an analysis of the schema in terms of inter-annotator

[1] The Preposition Project is a notable exception (cf. www.clres.com/prepositions.html).

Proceedings of the Fourth Linguistic Annotation Workshop, ACL 2010, pages 177–181,
Uppsala, Sweden, 15-16 July 2010.

agreement. It takes into account that the annotation schema is hierarchically ordered and allows for multiple annotations. Section 4 illustrates the application of the schema to the preposition *ohne* ('without') in German.

2 Properties of the Annotation Schema

There are no standardized features for an annotation of preposition senses in German. Our work is thus based on several reference works, which we analyzed and combined to develop the schema, namely *Duden Deutsch als Fremdsprache* (Duden, 2002) (a dictionary of German for foreign learners), *Deutsche Grammatik* from Helbig and Buscha (2001) (a grammar of German for foreign learners) (both usage-based), the *Lexikon Deutscher Präpositionen* (Schröder, 1986) (a dictionary of German prepositions) and an analysis of prepositions with a temporal meaning (Durell and Brée, 1993). Prima facie, the dictionary of German prepositions appears to be the most promising starting point because it includes a fine-grained feature-based analysis of preposition senses. However, it turns out that it is too complex for manual annotation, making use of more than 200 binary features to classify preposition meanings.

The annotation schema shows a hierarchically organized, tree-like structure. Beginning with a root node, types of preposition meanings branch to subtrees for different classes (e.g. local, temporal or causal) with differing depths or to individual, non-splitting branches (see Figure 1). For temporal and spatial interpretations, we use decision trees that help to guide the annotator through the annotation process.

Altogether the annotation schema includes the following list of top-level categories: SPATIAL, TEMPORAL, MODAL, CAUSAL, STATE, COMMUNALITY/COMMUTATIVE, TRANSGRESSION, AGENT, REDUCTION/EXTENSION, PARTICIPATION, SUBORDINATION, RECIPIENT, AFFILIATION, CORRELATION/INTERACTION, ORDER, THEME, SUBSTITUTE, EXCHANGE, COMPARISON, RESTRICTIVE, COPULATIVE, ADVERSATIVE, DISTRIBUTIVE, STATEMENT/OPINION, EXISTENCE/PRESENCE, CENTRE OF REFERENCE, and REALIZATION.

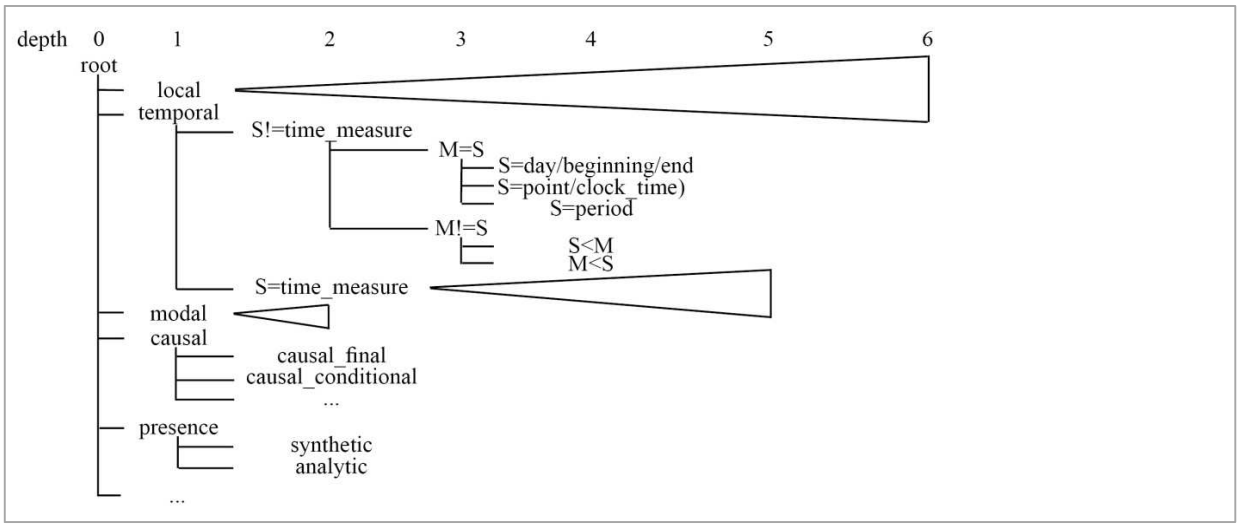


Figure 1: Hierarchical Annotation Schema

The schema allows cross-classification at every level. This is of particular importance for the classification of directional meanings. Directionality is introduced through cross-classification and not through copying the hierarchical structure of the local subtree.[2]

Another important property of the annotation schema is the possibility of multiple annotations for one preposition in context. For instance, a final distinction between a temporal and a causal interpretation cannot be drawn in example (3).

(3) *Feuer nach [temporal/causal] Blitzschlag*
'Fire after/because of lightning stroke'

In addition to the semantic categories, we use a feature 'governed' to label a preposition as governed by a lexical head whenever appropriate. Governed prepositions usually are assumed to be semantically empty but in some cases there is a discernible meaning for the preposition despite its being governed.

The preposition sense annotation is only one part of a bigger annotation project. Annotations on lexical (POS, morphology, countability, preposition meaning, noun meaning), syntactic (chunks), relational (internal and external dependencies), and

[2] During annotation, local and directional interpretations can be distinguished by case assignment in the majority of cases.

global (e.g. marking as a headline or part of a TV program in a newspaper, idiomaticity, telegraphic style) levels will serve as a basis for annotation mining to detect licensing conditions of PNCs.

3 An Analysis of Inter-Annotator Agreement in a Hierarchical Annotation Schema

A weighted kappa statistic (κ) forms a standard for assessing the feasibility of annotation schemata. Based on Cohen's seminal work (Cohen, 1968), Artstein and Poesio (2008) suggest the measure in (4), where κ is calculated as the weighted difference between observed and expected *disagreement.*

(4) $\kappa_w = 1 - \frac{D_o}{D_e}$

Two aspects of the present annotation schema prohibit a direct application of this statistic. 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 PRESENCE, while a second annotator makes use of the annotation ANALYTIC, 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, as in (4).

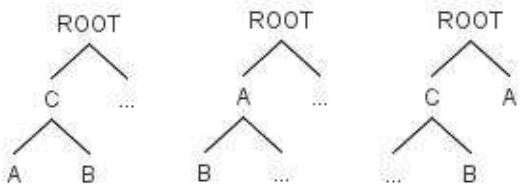
To deal with the first problem, the hierarchical structure of the annotation schema is included in the calculation of the weight coefficients for κ. Basically, two annotations are more closely related if either both annotations are dominated by the same set of nodes in the hierarchy, or one annotation is a direct subtype of the other one (as usual, we assume domination to be reflexive). 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.

As an illustration consider two senses A and B in the following configurations: a) A and B are directly dominated by C, a subtype of ROOT; b) A dominates B, A being a subtype of ROOT, and c) ROOT directly dominates A and C, and B is a subtype of C. Intuitively, c) is a case of clear disagreement, while in b) we find that one annotation is more specific than the other one, and in a), the annotators have at least agreed in a common supertype of the categories.

Consequently, the weight coefficient for disagreement should be highest in case c), but should be similar in cases a) and b).

(5)

a) b) c)



The weight coefficient is determined by the following formula, where $d_{k_j k_l}$ designates the depth of the lowest common dominating node of the two senses (and hence the cardinality of the set of dominating nodes minus 1).

(6) $$w_{k_j k_l} = \begin{cases} \frac{1}{2^{d_{k_j k_l}}}, & k_j \neq k_l \\ 0, & k_j = k_l \end{cases}$$

For the configuration a), the number of dominating nodes equals 2. Thus $d_{k_j k_l}$ equals 1, resulting in a weight coefficient of 0.5. For the configuration b), the cardinality of dominating nodes also equals 2, and again the weight coefficient is determined as 0.5. For c), however, the set of dominating nodes only contains ROOT, and consequently, the weight is determined as $1/2^0 = 1$.[3]

With regard to multiple annotations, we define new categories consisting of the combination of the used categories. To calculate the weight of disagreement between two combined categories, we compute the weights of all ordered pairs from the Cartesian product of the relevant categories and then calculate the arithmetic mean. As an illustration consider the following configuration: one annotator has assigned the senses A and B to a given preposition, where A and B are subtypes of C, while the second annotator has assigned B only. In this case, we determine the sum of disagreement between A and B and A and A, respectively, and divide it by the number of possible combinations (two in the present case). The following formula captures this idea.

(7) $w_{k_j k_l} = \frac{1}{|k_j||k_l|} \sum_{p \in k_j} \sum_{q \in k_l} w_{pq}$

Now, instead of determining the κ statistic on the basis of non-overlapping, i.e. mutually exclu-

[3] As we assume that dominance is reflexive, each supertype is a supertype of itself. Hence, the weights determined for the cases (5a) and (5b) are identical because A is a direct supertype of B. This would be different if A were an indirect supertype of B.

sive 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 examples of the prepositions *an, auf, bei, neben, unter,* and *vor.* The following table summarizes the results for the full set of sense annotations, for senses with subtypes *(local, temporal, causal, modal),* as well as for some individual senses.

Table 1: Subset of Weighted Kappa-values

subtree with the following root node	κ_w
ROOT	0.644
local	0.709
causal	0.575
modal	0.551
temporal	0.860
local_reference_plane	0.569
temporal_M=S_S=PERIOD	0.860

The overall result of 0.644 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, however, indicate the necessity to take a look at the data again and to identify sources of error.

4 Criteria for annotating *ohne* ('without')

The preposition ohne ('without') allows six different interpretations at top level, among them are the interpretations PRESENCE, COMITATIVE, and PARTICIPATION. The rule guided nature of the annotation schema will be illustrated by the following examples:

(8) *Die Anklage wirft dem ersten von*
The prosecution accuses the first of
drei Angeklagten, einem 32jährigen Mann
three accused a 32-year-old man
ohne Beruf, die Mitwirkung an
without profession the involvement at
allen drei Tötungsdelikten vor.
all three homicides PTKVZ
"The prosecution accuses the first of three defendants, a 32 years old man without a profession, of the involvement in all three homicides."

(9) *Ein mobiles Einsatzkommando überwältigte*
A mobile task force defeated
den Geiselnehmer, als er ohne das
the hostage-taker, when he without the
Kind den Gerichtssaal verließ.
child the court room left.
"A mobile task force defeated the hostage-taker, when he left the court room without the child."

(10) *Ein monetärer Schulterschluss ohne das*
A monetary closing of ranks without the
westliche Nachbarland wäre nicht
western neighboring country would be not
nur in Paris undenkbar.
only in Paris unthinkable.
"A monetary closing of ranks without involving the western neighbor would be unthinkable not only in Paris."

PARTICIPATION is defined as active or passive participation in an activity; COMITATIVE is defined as an abstract coactivity of two individuals or objects. PRESENCE, finally, characterizes the presence of an object or a property. With regard to *ohne,* the features have to be negated, i.e. denoting a *lack* of participation, co-activity, or *absence* of a feature. From the definition, it already follows that the external argument of a P with the interpretations PARTICIPATION or COMITATIVE is presumably event-like, but object-like with PRESENCE. COMITATIVE and PARTICIPATION, finally, are distinguished by the mutuality present in COMITATIVE, which is not present with PARTICIPATION, giving rise to an assignment of PRESENCE in (8), COMITATIVE in (9), and PARTICIPATION in (10).

5 Conclusion

We have presented an annotation schema for preposition senses in German that is based on usage-based grammars and dictionaries. It comprises a restricted set of less than 30 top level sense categories, and allows for multiple annotations of individual token if a maximal sense reduction cannot be achieved. The categories local, temporal, causal, modal and presence introduce hierarchical subtypes, access to the subtypes is partially guided by decision trees in the annotation process. The hierarchical structure of the annotation schema is also reflected in its validation in terms of inter-annotator agreement. Here, it became necessary to modify Cohen's κ to allow for overlapping categories and multiple annotations. The results reported here show that the schema is feasible for manual annotation of preposition senses.

References

Ron Artstein and Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34 (4): 555-596.

Timothy Baldwin et al. 2006. In search of a systematic treatment of determinerless PPs. In Patrick Saint-Dizier (ed.), *Syntax and Semantics of Prepositions*. Springer, Dordrecht, 163-179.

Christian Chiarcos, Stefanie Dipper, Michael Götze, Ulf Leser, Anke Lüdeling, Julia Ritz, and Manfred Stede. 2008. A flexible framework for integrating annotations from different tools and tagsets. *Traitement Automatique des Langues.* Special Issue Platforms for Natural Language Processing. ATALA, 49 (2).

Jacob Cohen. 1968. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological bulletin,* 70 (4): 213-220.

Florian Dömges, Tibor Kiss, Antje Müller and Claudia Roch. 2007. Measuring the Productivity of Determinerless PPs. *Proceedings of the ACL 2007 Workshop on Prepositions*, Prague*,* 31-37.

Duden. 2002. *Duden. Deutsch als Fremdsprache*. Bibliographisches Institut and F.A. Brockhaus AG, Mannheim.

Duden. 2005. *Duden. Die Grammatik*. Duden Band 4. Bibliographisches Institut & F.A. Brockhaus AG, Mannheim.

Martin Durell and David Brée. 1993. German temporal prepositions from an English perspective. In Cornelia Zelinsky-Wibbelt (ed.)*, The Semantics of Prepositions. From Mental Processing to Natural Language Processing*. De Gruyter, Berlin/New York, 295-325.

Gerhard Helbig and Joachim Buscha. 2001. *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht*. Leipzig, Langenscheidt.

Jochen Schröder. 1986. *Lexikon deutscher Präpositionen*. Leipzig, VEB Verlag Enzyklopädie.

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