A Pilot Study on the Semantic Classification of Two German Prepositions: Combining Monolingual and Multilingual Evidence

Simon Clematide
Computational Linguistics
University of Zurich
`simon.clematide@cl.uzh.ch`

Manfred Klenner
Computational Linguistics
University of Zurich
`klenner@cl.uzh.ch`

Abstract

This paper reports on the annotation and maximum-entropy modeling of the semantics of two German prepositions, *mit* ('with') and *auf* ('on'). 500 occurrences of each preposition were sampled from a treebank and annotated with syntacto-semantic classes by two annotators. The classification is guided by a perspective of information extraction, relies on linguistic tests and aims at the separation of semantically transparent and opaque meanings (that is of collocational constructions). Apart from descriptive statistical material, we present results of experiments using monolingual and multilingual evidence (the latter from informative English and Spanish translations) in order to predict the semantic classes.

1 Introduction

In linguistics, scientific grammars (Zifonun et al., 1997) as well as grammars for language learners (Helbig and Buscha, 2001) follow a long-standing tradition of semantic classification of prepositional phrases. However, it is less well-known which classification schemes can be used for automatic sense disambiguation, supporting for instance applications of information extraction and knowledge discovery.

In this pilot study, we want to gain experience of how to classify the semantic contributions of various prepositions from a multilingual perspective. Our main goal is to distinguish between semantically transparent contributions that prepositions can provide in a general or productive manner and the less transparent contributions in collocational constructions. Many prepositions are subcategorized by verbs (or adjectives) and the semantic contribution of a selected preposition is weak or unspecific—a fact that is often revealed by cross-lingual comparisons of subcategorization frames. In this study we want to assess the influence of syntactic dependencies and subcategorization on semantic classification. Therefore, we chose to take our material from a syntactically annotated treebank.

The rest of this paper is organized as follows. Section 2 presents related work and approaches. In Section 3, we describe our syntacto-semantic classification system used in the annotation of prepositions sampled from a German treebank. We also present the types of evidence used in the machine learning experiments for the automatic prediction of the classes. Section 4 contains a systematic evaluation of the performance of the different evidence that we have integrated in our approach.

2 Related Work

As Baldwin et al. (2009, p.134) have put it in their introduction to a special issue on that topic in the *Computational Linguistics Journal*: "Information extraction is one application where prepositions are uncontroversially crucial to system accuracy". The underlying task can be cast as preposition (word) sense disambiguation (WSD). It also has been recognized in the machine translation community that "prepositions are hard to translate" (Shilon et al., 2012, p.106). Although semantic information helps to tackle the translation task, the semantic class of a preposition does not perfectly determine the correct translation. As a consequence, these approaches do not strive to carry out preposition WSD, but to use semantic features in order to more directly map source prepositions to target prepositions (Li et al., 2005), be it rule-based (Agirre et al., 2009) or with machine learning given aligned bilingual data (Gustavii, 2005).

A great deal of work on preposition classification and WSD has been carried out on the English language. Most prominent the *Preposition*

Project (Litkowski and Hargraves, 2006) that uses a fine-grained classification scheme derived from the *Oxford Dictionary* (see also the *SemEval* Task on WSD of prepositions, Litkowski (2007)). Other elaborated classification schemes can be found as part of *VerbNet* (Kipper et al., 2004) and *PrepNet* (Saint-Dizier, 2008).

Annotated data is available from the *Penn Treebank* II (Marcus et al., 1994), where thematic roles occurring with prepositional phrases are marked, and *FrameNet* (Baker et al., 1998), which was annotated as part of the *Preposition Project*. There have been a couple of ACL-SIGSEM workshops on prepositions (the last one in 2007) covering all aspects of preposition processing (not only the semantics).

On the methodological side, preposition disambiguation sometimes is coupled with semantic role resources, e.g. O'Hara and Wiebe (2009). There, traditional features for WSD (e.g. the preposition, stem of embedded noun, POS and stem of words in a fixed window around the preposition) are augmented with semantic features stemming from knowledge resources such as FrameNet and *WordNet* (Fellbaum, 1998). In O'Hara and Wiebe (2009), a new feature, hypernym collocation (the WordNet hypernym of the embedded noun), is used to carry out disambiguation relative to either coarse-grained Penn treebank functional roles or more sophisticated FrameNet roles. Syntactic information, e.g. the syntactic function of the PP, is ignored in their system (in contrast to our approach).

As can be seen from the discussion above, there is no canonical classification scheme for preposition disambiguation. Furthermore, the semantic class that a preposition can take is language specific. For German, there are but a few approaches (Hartrumpf et al., 2006; Müller et al., 2011). Müller et al. (2011) rely on an annotation scheme derived from various traditional linguistic theories. 22 prepositions are modeled on the basis of 27 top-level senses. A sense hierarchy is defined (especially for temporal and spatial senses) in order to allow for a more flexible and fine-grained classification. Manually specified decision trees are then used to produce the gold standard classifications.

This scheme is, for our purposes, far too fine-grained and also hard to automatically model by machine learning. However, if their resources

sem\syn	opp	mod	vmod	?mod	–	p	$\sum$
verbal	131	2	3		3		139
nominal	2	120	2	2	3		129
coll	42	3	8			2	55
TEM			6				6
MOD	3		8	4	1		16
LOC	8	10	77	8	5	2	110
DIR	16		7				23
TLOC			9				9
CAU			3	3	1		7
?		1	2		2	1	6
$\sum$	202	136	125	17	15	5	500

Table 1: Distribution of semantic functions of *auf* (*on*) in relation to the syntactic function. The syntactic function "predicative" is labelled as "p".

sem\syn	vmod	mod	opp	?mod	–	$\sum$
verbal	16	8	107	2	1	134
nominal	4	53	7			64
coll	3		1			4
TEM	4			1		5
MOD	46	2	2	6	4	60
INS	75	3	4	1	1	84
ORN	5	56		4	1	66
COM	30	10	3	2	1	46
IDE	8	1		7		16
SIZ	4	6	1	1		12
?	1	3		1	3	8
$\sum$	196	142	125	25	11	499

Table 2: Distribution of the semantic function of *mit* in relation to the syntactic function. The syntactic function predicative is not shown in the table because it appeared only once.

were available, we could probably map their scheme to our scheme. No attempt was made by Müller et al. (2011) to learn a model for preposition classification based on their semantic classes. Their approach based on logistic regression as described in Kiss et al. (2010) focuses on determiner omission in PPs.

The work of Hartrumpf et al. (2006) is geared towards a semantic formalism called *MultiNet* (Helbig, 2006), it fully relies on this proprietary resource.

3 Methods

3.1 Resources

As mentioned in Section 2, the Penn Treebank comprises shallow semantic annotations to prepositional phrases (PP). There, a distinction is made between six semantic classes of PPs (and, thus, prepositions): locative, direction, manner, pur-

pose, temporal, and extent. Unfortunately, none of the large German treebanks (TIGER (Brants and Hansen, 2002), Tüba-D/Z (Telljohann et al., 2004)) provide such a comparable rudimentary scheme that could be a starting point for our pilot study. There is no resource, we could use, although one is currently being developed by another group (Müller et al., 2011), but it is not yet released. Since we believe that treebanks could benefit from such an additional annotation layer, we decided to work with a German treebank, the *Tübinger Baumbank* Tüba-D/Z 7.0. It comprises about 65,000 annotated sentences, besides phrase structure, also topological fields and grammatical functions are specified. PPs can act as obligatory or optional (opp) complements of verbs, or as adjuncts (vmod). In the current study, we mainly focus on PPs acting as verb complements (opp) or adjuncts (vmod).

From the ten most frequent prepositions in the Tüba-D/Z we have chosen one with a predominant local and temporal meaning (*auf* 'on') and one with more broader meaning spectrum (*mit* 'with'). We randomly sampled 500 occurrences of each preposition from the Tüba-D/Z and annotated each preposition according to our classification scheme described below.

3.1.1 Semantics of *auf* and *mit*

The intended application is information extraction and question answering. Accordingly, our semantic classes had to be tightly coupled with question words. That is, the way users may ask, determines the granularity of the classification scheme. Typical interrogative words and phrases are *how* (modal), *how long* (temporal, duration), *when* (temporal, time point), *where* (locative).

In the case of *auf* (cf. Table 1), we distinguish between locative (LOC *where*), directional (DIR *where to*), temporal (TEM *when*, *how long*), modal (MOD *how*), and causal (CAUS *why*) PPs. If in a temporal PP the noun is an event (e.g. *party*), then often a locative or a temporal reading is possible (e.g. *when or where did he laugh? - at his party*). We use TLOC to refer to this usage. If the PP acts as a modifier of an adjective or noun, it is annotated with "nominal" (e.g. 'the cup *on the table*'). For the preposition *auf*, we have annotated currently only adjuncts and verb complements with their semantic classes. In case that the verb governs an otherwise semantically vacuous preposition (*warten auf* 'to wait for'),

the preposition is marked with "verbal". Finally, any idiomatic expression comprising a PP having a non-compositional meaning like *auf den Putz hauen*, 'to kick up one's heels' is annotated as collocational ("coll"). The preposition does not contribute any semantics in these cases. Sometimes no decision was possible (e.g. given sentence fragments, missing global context, unclear semantics), we used "?" to annotate these cases.

Table 1 shows the distribution of these classes and their syntactic realization. Verb/preposition collocations form the largest class (139), followed by nominal modification (129) and locatives (110). Syntactically, there are three groups to be distinguished: PP complements (opp, 202), NP and PP modification (mod, 136) and adjuncts (vmod, 125). The table reveals a moderate number of interpretation divergences between the Tüba-D/Z annotators and us. Some stem from structural ambiguity (e.g. "?mod" denotes PP attachment ambiguities), and are to be expected. Ideally, however, if a PP bears the functional label "mod", it should be classified as "nominal" in our scheme. Also, a "vmod" should not be annotated as "verbal", since "verbal" means that the preposition is vacuous, while "vmod" means that it acts as an adjunct. For instance, we disagreed with 3 "vmod" (adjuncts) and interpreted them as verb-preposition collocations, also 2 "vmod" are better classified as "nominal" in our view. However, the majority of decisions does not contradict or even is in line with the functional assignments of the Tüba-D/Z. For example, of the 136 "mod" (NP or PP modifications), we placed 120 in our corresponding class "nominal".

In the case of *mit* (cf. Table 2), the syntactic classification labels "verb", "nominal" and "coll" are used as introduced above for *auf*. The prepositions *auf* and *mit* also share two core semantic classes, namely TEM (temporal) and MOD (modal). The other semantic classes of *mit* are: COM for comitative use (*to watch a movie with a friend*), ORN for ornative use (*to tell with humor*), SIZ indicating size or extent (*to demonstrate with 100 people against*), INS for the instrument reading which is a subclass of MOD (modal) (*to break with a hammer*), and IDE for identity (*with him, hope enters the room* meaning: *he represents/is identical with hope*).

As with *auf*, there are some divergences between the functional annotations of the Tüba-D/Z

and our annotation decisions, especially concerning "vmod" and "mod". We have not fully traced these divergences back to their origins, but see the previous discussion in the context of *auf*.

3.1.2 Inter-Annotator Agreement

We have measured inter-annotator agreement in two stages: after our initial annotation round, and after some discussion and refinements of our annotation scheme in a second step on a harmonized version of the data. One reason for disagreement concerning *mit* was the annotation with "ornative": a rather sophisticated annotation scheme would allow the use of ORN even in cases where it is modal, but also implicitly qualifies the subject of the sentence (*he says it with a gentle voice*). In these cases, however, it is more natural to ask *how* (*has he said it*), so we disallowed ORN in such examples.

We report the annotator agreement as percentage of agreeing pairs and as Cohen's κ. The initial inter-annotator agreement for *mit* was 85% (κ = .82), while after harmonization it was 91.8% (κ = .90) and 92% (κ = .90) between the harmonized version and the separately created initial annotations of the two annotators, respectively. With *auf* the agreement was lower, namely initially 74% (κ = .67). After harmonization is was 84.8% (κ = .81) and 86.2% (κ = .83), respectively. The main source of confusion here was the annotation scheme of PPs in the context of local verbs (LOC and DIR). The question was whether to treat these roles as adjuncts or as verb complements. Also the decision when to treat a verb-preposition combination as a collocation or not, was not sufficiently well described and operationalized in the guidelines.

The rationale behind our two-stage procedure was to first independently create annotation strategies based on existing classes from the literature and to later refine them to valid annotation guidelines based on the evidence found in the data.

3.1.3 Multilingual Evidence

As already mentioned, prepositional semantics is language-specific: The semantic classes a preposition might express do vary between languages, the semantic contributions given by a preposition in one language are often realized by different prepositions in different languages. Moreover, the semantic functions a preposition (e.g. *mit*) and its direct translation ('with') can bear, might differ.

The identity reading of German *mit* is not possible for English 'with'.

The question is, whether a multilingual perspective (for instance in the form of Statistical Machine Translation (SMT)) helps determining the semantic class of a given preposition in the source language. Tables 3 and 4 give a detailed overview of how the prepositions *mit* and *auf* are translated into English and Spanish by *Google Translate*.[1] For instance, *mit* is translated into English as *with*, *of*, *to*, *by*, *in*, or not at all ("0"). Of course, there are predominant translations, for instance *mit* was translated 372 times by *with* and *con*. There is also a tendency to choose equivalent prepositions across languages, e.g. *a* and *to* (Table 4: 71 cases), but quite often different prepositions are selected. Since we use imperfect translations from SMT we cannot be sure whether the aforementioned differences stem from mistranslations or whether they reveal a true difference. In order to clarify this question one could exploit parallel treebanks. However, currently available resources covering German such as SMULTRON (Volk et al., 2010) still have a limited size (approx. 2500 sentences).

The question is whether inter-language divergence of preposition usage helps to determine the semantic class of a preposition in the source language. Or more technically, whether there is a correlation between semantic classes of the source language preposition and a translation made by SMT. Even if such a correlation turns out not to be a strong one, it might nevertheless help as a feature in a machine learning model.

3.1.4 Annotation and Translation: Examples

For illustration purposes, we give two examples of semantic annotations of PPs and the mapping of the German prepositions therein to English prepositions via automatic translation with Google translate.

In the first case, *auf* does not carry any semantics, it is part of the verb (*warten auf*). Accordingly, it is annotated as "verbal". In English, the corresponding verb construction is 'to listen to',

---

[1] For this experiment, we manually mapped the prepositions from the translated sentences using the phrase alignment visualization of `http://translate.google.com`. English and Spanish was chosen since according to `http://matrix.statmt.org/matrix` the translation quality of German to English and Spanish is best and at the same time both target languages belong to different language families.

es\en	with	0	by	to	of	in	on	about	as	from	$\sum$
con	372	7	1	6	1			1			388
0	3	48		2	1	1	1		1		57
de	10	5	1		6					1	23
a	7	5	1		1						14
por	1	1	6								8
en	1					2	2				5
como		2									2
y		2									2
para				1							1
$\sum$	394	70	9	9	8	4	3	1	1	1	500

Table 3: Translations (Google Translate) of German *mit* in English and Spanish. Columns and rows are ordered by margin frequencies.

es\en	on	to	0	in	at	of	for	about	by	with	around	$\sum$
en	182	7	7	27	17		1		1		2	244
a	8	71	6	2	10	2	2			1		102
0	7	7	33			3	6	1				57
de	9	10	2	5	1	17	3					47
sobre	15	2						1				18
para		8					4					12
por	5		1				1		1			8
con	1	2			1					1		5
contra	2											2
$\sum$	229	107	49	34	29	22	17	2	2	2	2	495

Table 4: Translations of German *auf* in English and Spanish. Translations appearing only once are not shown.

which is correctly identified by Google Translate. The sentence pairs are: *Man muss auf diesen Aufschrei hören* and 'You have to listen to this outcry'.

The second examples illustrates that the same semantic class, LOC (local), might be realized by two different prepositions in German and in English. The preposition *auf* in German can be used to indicate the 'place of living' of a person, if it is a small island (like Sardinia). This is not possible in English. The sentence pairs are: *Selbst wenn sie in entlegenen Städtchen auf Sardinien leben* and 'Even if they live in remote town in Sardinia'.

Note that in these examples *auf* was not mapped to its direct translation which is 'on'.

3.2 Supervised Machine Learning Approach

In order to measure the difficulty of an automatic classification of the syntacto-semantic classes expressed by *auf* and *mit* we conducted several experiments with the Maximum-Entropy Modeling tool MegaM (Daumé III, 2008).[2] For this pilot study, we focused on simple features gained from

the syntactical configuration (perfect data from the Tüba-D/Z), textual data from the context, and multilingual evidence from Spanish and English translations (imperfect Google translations).

In Section 4 we present and analyze the contribution of the following feature sets:

head Word, part of speech (POS), and lemma of the head word (typically a noun) of the dependent phrase of the preposition, for instance, the head of *mit Sorgfalt* is *Sorgfalt* 'care'. In case of coordinated PPs and multi-word heads, the first token was selected.

syntax The syntactic function of the PP taken from the TübaD/Z.

neighbor Word, POS, and lemma of the preceding and following token.

context Word, POS, and lemma in a window of 5 preceding and following tokens (taken as a bag of words, lemmas and POS).

en English translation of the preposition produced by the Google translation of the German sentence.

es Spanish Google translation of the preposition.

---

[2]Maximum-Entropy modeling is also known as logistic regression. In our experiments, we used the default regularization parameter $\lambda = 1$ of MegaM.

4 Results and Discussion

The evaluations described below assess the performance improvement for the multi-class predictions of our annotated prepositions (500 occurrences each) by using different sets of features as evidence. We evaluate against a baseline system which basically predicts the majority class given the lack of any additional evidence. All results are reported as mean accuracy computed by cross-validation. No stratification of class labels has been applied to the folds of the cross-validation. Accuracy is the proportion of true classifications delivered by the system.

4.1 Syntacto-Semantic Classification

We performed a 10-fold cross-validation evaluation for the scenario of predicting the full set of all syntactic and semantic classes (cf. Table 1 and 2).

The evaluation results of *auf* are shown in Table 5. The best system uses the feature sets "head", "neighbor" and "syntax", however, "syntax" is by far the strongest feature. If perfect syntactic analyses are not available, "head" and "neighbor" information can compensate for more than 2/3 of the performance gain. The effect of "syntax" is especially strong for *auf* because the nominal modifiers are classified according to syntactic criteria only. A future, more semantically oriented classification of noun modifiers will probably weaken this effect. Multilingual evidence from informative Google translations improves considerably over the (weak) baseline. Combining the evidence from Spanish and English performs slightly better than each language separately does. Therefore, translations into multiple languages are useful for the case of *auf*. However, the best systems are those without any translation evidence from Spanish or English.

Table 6 shows the corresponding results for *mit*. The overall performance is lower but the feature sets have a very similar ranking of predictive power. The lower performance stems from the fact that *mit* has 11 syntacto-semantic classes with a more uniform distribution than *auf* (10 classes). The best system without the feature "syntax" involves 3 different feature sets, "context", "neighbor" and "en". However, these feature sets can only compensate for less than half of the performance gain of the feature set "syntax" derived from the treebank syntax structure. The best system performance is reached if either English or

Evidence	Mean	SD	$\triangle \text{abs}_{bs}$	$\triangle \text{rel}_{bs}$
baseline	25.4	7.5		
head	27.2	7.8	+1.8	+7.1
en	38.6	10.5	+13.2	+52.0
es	39.0	8.7	+13.6	+53.5
context	45.4	9.9	+20.0	+78.7
neighbor	53.4	7.6	+28.0	+110.2
syntax	68.6	7.2	+43.2	+170.1
en/es	39.4	8.3	+14.0	+55.1
head/neighbor	58.2	6.4	+32.8	+129.1
head/syntax/neigh.	**71.0**	6.6	+45.6	+179.5

Table 5: Performance of feature sets for syntacto-semantic classification accuracy: *auf* ($N = 500$). The column "Mean" contains the average accuracy computed from the cross-validation sets. The column $\triangle \text{abs}_{bs}$ contains the absolute performance gain with respect to the baseline. $\triangle \text{rel}_{bs}$ expresses the relative performance gain. The last row contains the feature set with the best performance.

Evidence	Mean	SD	$\triangle \text{abs}_{bs}$	$\triangle \text{rel}_{bs}$
baseline	26.8	7.1		
head	28.8	7.1	+2.0	+7.5
context	34.6	5.8	+7.8	+29.1
neighbor	36.2	4.0	+9.4	+35.1
syntax	46.4	8.4	+19.6	+73.1
neighbor/context/en	40.4	6.7	+13.6	+50.7
head/syn./neigh.	57.2	7.5	+30.4	+113.4
head/syn./neigh./en	**57.4**	8.2	+30.6	+114.2
syn./neigh./cont./es	**57.4**	7.9	+30.6	+114.2

Table 6: Performance of feature sets for syntacto-semantic classification accuracy: *mit* ($N = 500$).

Spanish evidence is added. However, the improvement given by multilingual evidence is rather small.

4.2 Semantic Classification

In a further evaluation, we measured how well the purely semantic classes (i.e. those without "nominal", "verb" and "coll") can be predicted. For *auf* we have 171 cases with a defined semantic classification, for *mit* 290. Due to the smaller training sizes we performed 5-fold cross-validation.

Table 7 illustrates the problems from the skewed distribution of semantic classes in the case of *auf*: Just guessing the largest class LOC represents a baseline decision which is hard to beat. Only the feature set "head" can improve over this baseline, all other features either deteriorate the system performance or do not improve it. Interestingly, the best system combines the translation evidence from Spanish with the feature set "head". Adding

Evidence	Mean	SD	$\triangle\text{abs}_{bs}$	$\triangle\text{rel}_{bs}$
baseline	72.9	6.7		
head	75.3	6.4	+2.4	+3.2
head/syntax/neigh./es	**77.6**	7.7	+4.7	+6.5
head/es	**77.6**	7.7	+4.7	+6.5

Table 7: Performance of feature sets for semantic classification accuracy: *auf* ($N = 171$). The following classes are considered: LOC, DIR, MOD, TLOC, CAU, TEM.

Evidence	Mean	SD	$\triangle\text{abs}_{bs}$	$\triangle\text{rel}_{bs}$
baseline	26.2	9.9		
head	27.6	8.8	+1.4	+5.3
context	36.6	12.3	+10.3	+39.5
neighbor	39.3	13.7	+13.1	+50.0
syntax	42.1	4.5	+15.9	+60.5
head/neigh./en/es	40.7	11.3	+14.5	+55.3
head/syntax/neigh.	**52.4**	5.7	+26.2	+100.0

Table 8: Performance of feature sets for semantic classification accuracy: *mit* ($N = 290$). The following classes are considered: TEM, MOD, INS, ORN, COM, IDE, SIZ.

more feature sets does not improve the results (see Table 7 second last row).

The less skewed distribution of semantic classes in the case of *mit* allows for a significant improvement over the baseline system. Table 8 shows that most feature sets have a beneficial effect, and therefore, classification performance is almost doubled by the best system. In contrast to the syntacto-semantic classification, multilingual evidence does not contribute to the best system. The only configuration where multilingual evidence improves performance appears if the syntactic dependency information from the treebank is dropped. The best system without the feature set "syntax" relies on English and Spanish evidence.

The results of our experiments in using multilingual evidence for the syntacto-semantic and semantic classification of prepositions are mixed. The syntactico-semantic classification of *auf* works best without multilingual evidence although there is a weak correlation between the feature sets "en" and "es" and the syntacto-semantic classes. However, the best system of the syntactico-semantic classification of *mit* profits from added multilingual evidence although this evidence taken as a single feature set cannot beat the baseline.

For the purely semantic classification, no improvement over the baseline can be found by the

multilingual evidence for both prepositions. Still, multilingual evidence helps in these cases where syntactic information is not valuable (in the case of *auf*), or if we mute this feature (in the case of *mit*).

5 Conclusion

Our annotation and modeling experiments illustrate the different semantic and distributional characteristics of the considered German prepositions *auf* and *mit*. The skewed distribution of the semantic classes of *auf* represent a challenge for any classifier. If small semantic classes should be detected, more training material is needed for these cases. The application of Active Learning techniques (Settles, 2012) might help to efficiently collect such data.

Our experiments with maximum entropy modeling indicate that informative Google translations of prepositions do not lead to a significant performance improvement in semantic classification. Simple monolingual contextual features generally perform better. The inclusion of perfect (i.e. treebank-derived) syntactic dependency information generally performs best. However, for practical systems only imperfect syntax analyses from error-producing parsers are available. Future research is needed to assess the performance decrease if parser output is provided instead of hand-crafted manual annotation.

Another topic for future work is the integration of further language resources. Bilingual lexicons such as dict.cc[3] contain information about semantically void subcategorized prepositions, for instance *auf jdn warten* is linked to *to wait for sb*. Statistical collocation analyses derived from large German corpora are provided by services such as "Wortschatz Leipzig"[4] or "Digitales Wörterbuch der Deutschen Sprache"[5].

Given the available amount of electronic texts, the application of distributional semantics for preposition disambiguation and for modeling of the semantic fingerprint of prepositions also seems promising (cf. (de Cruys and Apidianaki, 2011)).

Finally, contextual features might profit from synonym expansion or synonym set classification, a technique also used by Kiss et al. (2010).

---

[3]See `http://www.dict.cc`
[4]See `http://wortschatz.uni-leipzig.de`
[5]See `http://dwds.de`

References

Eneko Agirre, Aitziber Atutxa, Gorka Labak, Mikel Lersundi, Aingeru Mayor, and Kepa Sarasola. 2009. Use of rich linguistic information to translate prepositions and grammar cases to Basque. In *Proceedings of the XIII Conference of the European Association for Machine Translation (EAMT)*, pages 58–65.

Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. In Christian Boitet and Pete Whitelock, editors, *COLING-ACL*, pages 86–90. Morgan Kaufmann Publishers / ACL.

Timothy Baldwin, Valia Kordoni, and Aline Villavicencio. 2009. Prepositions in applications: A survey and introduction to the special issue. *Computational Linguistics*, 35(2):119–149.

Sabine Brants and Silvia Hansen. 2002. Developments in the TIGER annotation scheme and their realization in the corpus. In *Proceedings of the Third Conference on Language Resources and Evaluation (LREC 2002)*, pages 1643–1649, Las Palmas.

Hal Daumé III. 2008. MegaM: Maximum entropy model optimization package. ACL Data and Code Repository, ADCR2008C003.

Tim Van de Cruys and Marianna Apidianaki. 2011. Latent semantic word sense induction and disambiguation. In Dekang Lin, Yuji Matsumoto, and Rada Mihalcea, editors, *ACL*, pages 1476–1485. The Association for Computer Linguistics.

Christiane Fellbaum. 1998. *WordNet: An Electronic Lexical Database*. Bradford Books.

Ebba Gustavii. 2005. Target language preposition selection – an experiment with transformation-based learning and aligned bilingual data. In *Proceedings of the XIII Conference of the European Association for Machine Translation (EAMT)*, pages 112–118.

Sven Hartrumpf, Hermann Helbig, and Rainer Osswald. 2006. Semantic interpretation of prepositions for NLP applications. In *Proceedings of the Third ACL-SIGSEM Workshop on Prepositions*, pages 29–36.

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

Hermann Helbig. 2006. *Knowledge Representation and the Semantics of Natural Language*. Cognitive Technologies. Springer.

Karin Kipper, Benjamin Snyder, and Martha Palmer. 2004. Using prepositions to extend a verb lexicon. In *Proceedings of the HLT/NAACL Workshop on Computational Lexical Semantics*, pages 23–29.

Tibor Kiss, Katja Keßelmeier, Antje Müller, Claudia Roch, Tobias Stadtfeld, and Jan Strunk. 2010. A logistic regression model of determiner omission in PPs. In Chu-Ren Huang and Dan Jurafsky, editors, *COLING (Posters)*, pages 561–569. Chinese Information Processing Society of China.

Hui Li, Nathalie Japkowicz, and Caroline Barrière. 2005. English to Chinese translation of prepositions. In *Canadian Conference on AI*, volume 3501 of *Lecture Notes in Computer Science*, pages 412–416. Springer.

Kenneth C. Litkowski and Orin Hargraves. 2006. Coverage and inheritance in The Preposition Project. In *Third ACL-SIGSEM Workshop on Prepositions*, pages 37–44.

Kenneth C. Litkowski. 2007. CLR: Integration of FrameNet in a text representation system. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 113–116, Prague, Czech Republic, June. ACL.

Mitchell P. Marcus, Grace Kim, Mary Ann Marcinkiewicz, Robert MacIntyre, Ann Bies, Mark Ferguson, Karen Katz, and Britta Schasberger. 1994. The Penn treebank: Annotating predicate argument structure. In *HLT*, pages 114–119. Morgan Kaufmann.

Antje Müller, Claudia Roch, Tobias Stadtfeld, and Tibor Kiss. 2011. Annotating spatial interpretations of German prepositions. In *ICSC*, pages 459–466. IEEE.

Tom O'Hara and Janyce Wiebe. 2009. Exploiting semantic role resources for preposition disambiguation. *Computational Linguistics*, 35(2):151–184.

Patrick Saint-Dizier. 2008. Syntactic and semantic frames in PrepNet. In *IJCNLP*, pages 763–768. ACL.

Burr Settles. 2012. *Active Learning*, volume 6 of *Synthesis Lectures on Artificial Intelligence and Machine Learning*. Morgan & Claypool.

Reshef Shilon, Hanna Fadida, and Shuly Wintner. 2012. Incorporating linguistic knowledge in statistical machine translation: Translating prepositions. In *Proceedings of the EACL-2012 Workshop on Innovative Hybrid Approaches to the Processing of Textual Data*, pages 106–114.

Heike Telljohann, Erhard W. Hinrichs, and Sandra Kübler. 2004. The Tüba-D/Z Treebank: Annotating German with a Context-Free Backbone. In *LREC*, pages 2229–2232. European Language Resources Association.

Martin Volk, Anne Göhring, Torsten Marek, and Yvonne Samuelsson. 2010. SMULTRON (version 3.0) — The Stockholm MULtilingual parallel TReebank. electronic.

Gisela Zifonun, Ludger Hoffmann, and Bruno Strecker. 1997. *Grammatik der deutschen Sprache*. Schriften des Instituts für deutsche Sprache; 7.1-3. de Gruyter, Berlin; New York.