Distributional regularities of verbs and verbal adjectives: Treebank evidence and broader implications

Daniël de Kok and **Patricia Fischer** and **Corina Dima** and **Erhard Hinrichs**
Department of General and Computational Linguistics, University of Tübingen
{daniel.de-kok, patricia.fischer, corina.dima, erhard.hinrichs}
@uni-tuebingen.de

Abstract

Word formation processes such as derivation and compounding yield realizations of lexical roots in different parts of speech and in different syntactic environments. Using verbal adjectives as a case study and treebanks of Dutch and German as data sources, similarities and divergences in syntactic distributions across different realizations of lexical roots are examined and the implications for computational modeling and for treebank construction are discussed.

1 Introduction

Due to processes of word formation such as derivation and compounding, lexical roots can be realized in different parts of speech and in different syntactic environments. For example, the derivational suffix *-able* can turn the verbal root *derive* in English into the adjective *derivable*, and the derivational suffix *-ity* can turn *derivable* into the noun *derivability*. A direct corollary of this polycategorial property of lexical roots and their morphological derivatives is their participation in different syntactic constructions and contexts, each of which comes with their construction-specific frequency distributions of collocations, syntactic arguments, modifiers, and specifiers.

In structuralist theories of language, the characterization of linguistic categories and structures in terms of their distributional behavior provides the key insight underlying distributional accounts of phonology, morphology, and syntax, most famously articulated by Harris (1951) and of semantics, as proposed by Firth (1957). The correct modeling of the interface of derivational morphology and syntactic derivations was also one of the central issues in the early days of generative grammar, with proponents of Generative Semantics (Lees, 1960) arguing for a transformational, syntactic account of word formation and Chomsky (1970) arguing for a non-transformational, interpretative account. In non-derivational, lexicalist theories of grammar such as Head-Driven Phrase Structure Grammar, the sharing of argument structure for lexical roots realized in different word classes is modeled by the non-transformational mechanism of lexical rules and sharing of valence information (see Gerdemann (1994) for such an account for nominalizations in German). Most recently, distributional theories of natural language have also served as an inspiration for distributional modeling of words as word embeddings in computational linguistics (Mikolov et al., 2013; Pennington et al., 2014).

Linguistically annotated corpora, so-called *treebanks*, offer excellent empirical resources for the study of the realization of lexical roots in different morpho-syntactic categories and constructions, provided that their annotations are rich enough to capture relevant information about derivational morphology and lemmatization.

2 Case Study

The purpose of the present paper is to systematically study similarities and divergences in syntactic distributions across different realizations of lexical roots. In particular, we are interested in finding out if the syntactic distribution of a particular realization of a lexical root can serve as an additional information source in modeling the meaning of other, possibly less frequent realizations of the same lexical root.

Proceedings of the 16th International Workshop on Treebanks and Linguistic Theories (TLT16), pages 1–9,
Prague, Czech Republic, January 23–24, 2018.

Keywords: treebank, annotations, verb, adjective, comparison, Dutch, German, PP, preposition

The paper focuses on a case study of the morpho-syntactic category of adjectives, and within that category on verbal adjectives such as *gegeten* 'eaten' in Dutch and *verloren* 'lost' in German, which are derived from the verbal roots *eten* 'to-eat' and *verlieren* 'to-loose', respectively. Verbal adjectives are of primary interest here since their syntactic distribution is that of an adjective, yet at the same time resembles the syntactic distribution of the verbs from which they are derived. As other adjectives, verbal adjectives occur in three syntactic environments: in attributive, pre-nominal position, in predicative position and in adverbial position, as exemplified in (1a), (1b), and (1c) respectively.

(1) a. [Die [**gewählten / wählenden**] / Weitere [**gewählte / wählende**]] Mitglieder stimmten zu .
[the [elected / voting] / more [elected / voting]] members agreed .

b. Die Mitglieder sind **gewählt** .
The members are elected .

c. Sie gaben **frustriert** auf .
They gave frustrated in .

Such adjectives are identical in form to the past participles of the verbs they are derived from. Their adjectival nature is underscored by the fact that they exhibit the same strong/weak inflectional alternation characteristic of adjectives in attributive position, as shown in (1a). Such inflectional variation does not occur in predicative and adverbial position so that the distinction between past participle verbs and verbal adjectives cannot be established in terms of linguistic form, but only in terms of syntactic environment. Moreover, present participles occur as predicative adjectives only in lexicalized cases (Lenz, 1993).

At the same time, verbal adjectives share the same type of arguments and modifiers with the verbs that they derive from. This includes in particular prepositional arguments and modifiers. Since the correct attachment of prepositional phrases is notoriously difficult for rule-based and statistical parsers alike, the present study focuses on the distributions of prepositions that are governed by verbs and verbal adjectives. We focus on prepositions in PP modifiers, as well as prepositional complements (PC) of verbs, as illustrated in (2).

(2)

		PP					PC		
Die	im	Deutschland	**gekauften**	Fahrräder	sind	gegen	Diebstahl	**versichert**	.
The	in	Germany	bought	bikes	are	against	theft	insured	.

As discussed in more detail in Section 4, our goal is to predict the distribution of prepositions governed by verbal adjectives from the distributions of the corresponding verbs. When dealing with ambiguous PP attachments to verbal adjectives, the information gained from the distribution of the corresponding verbs can be instrumental in choosing the correct attachment, especially in the case of predicative adjectives.

The current study uses data from two treebanks: the Lassy Large treebank (Van Noord et al., 2013) of written Dutch and the TüBa-D/DP treebank of written German (taz/Wikipedia sections).

3 Delineating the Domain of Verbal Adjectives

Since verbal adjectives combine properties of verbs and adjectives, it is to be expected that there are certain cases where the boundaries between verbal adjectives and verbs/adjectives are not as clear. In this section, we discuss these boundaries and their ramifications for our study.

3.1 Distinguishing Verbal Adjectives from Verbal Participles

An ongoing topic of debate is the word category of past participles that are governed by verbs which can either be auxiliary or copular. Consider (3), where the Dutch past participle form *gewaarborgd* 'guaranteed' can be analyzed as a verb participle that forms the verb cluster governed by the auxiliary verb *zijn* 'are' or a verbal adjective that is the predicative complement to the copular verb *zijn*.

(3) De obligaties [zijn / worden] **gewaarborgd** door het Vlaams Gewest .
The bonds [are / are-being] guaranteed by the Flemish region .

In Dutch, such ambiguities occur with several verbs that can have auxiliary and copular readings, most prominently *zijn* 'to-be', *worden* 'to-become', and *blijven* 'to-remain'.[1] In German only past participles

[1] The ambiguity does not occur in all word orders (Zwart, 2011).

governed by the verb *sein* 'to-be' (the so-called Zustandspassiv) are considered ambiguous. For the present work, we simply treat such participles as ambiguous and evaluate them as a separate set, as described in Section 4.[2]

3.2 Deverbal Adjectives

Although verbal adjectives can be derived productively, they can undergo various degrees of lexicalization, which can result in changes in argument structure or semantics as consequences. We will refer to such adjectives as *deverbal adjectives*, and we use the term *verb-derived adjective* throughout this paper as a cover term for verbal and deverbal adjectives. Deverbal adjectives pose two interesting challenges for the present study: First, they can give rise to new senses of a surface form, along with corresponding shifts in distributions of prepositions. For example, the German adjective *geschlossen* in *geschlossene Gesellschaft* 'closed society' has diverged in meaning from the participle of the verb *schließen* (*geschlossen*). However, it is also possible to use *geschlossen* in its verbal sense such as in *geschlossene Tür* 'closed door'. These two senses are combined with different prepositions. For example, *die durch Klaus geschlossene Tür* 'the by Klaus closed door' is a plausible PP-modification, while *die durch Klaus geschlossene Gesellschaft* is not. Unfortunately, this problem cannot be solved without word sense disambiguation, which (paradoxically) relies on co-occurrence statistics. Consequently, in such cases we model the preposition distribution of all senses together.

Secondly, some forms have transformed morphologically and syntactically into full adjectives, while retaining co-occurrence preferences. For example, the Dutch adjective *onomkeerbare* 'irreversible' in (4a) derives from the verb *omkeren* 'to reverse'. The adjective *onomkeerbaar* still accepts the same PP modifier *wegens klimaatverandering* 'by climate-change' as the past participle *omgekeerd* 'reversed' (4b). As discussed in Section 4, we include such adjectives in our German data set tracing them back to their original verb lemma where possible.

(4) a. ... het wegens klimaatverandering **onomkeerbare** process van zeespiegelstijging ...
... the because-of climate-change irreversible process of sea-level-rise ...

b. Het process van zeespiegelstijging kan wegens klimaatverandering niet **omgekeerd** worden .
The process of sea-level-rise can because-of climate-change not reversed become .

4 Empirical Basis

To study the distribution of prepositions governed by verbs and verbal adjectives, we extract co-occurrences between (i) prepositions; and (ii) verbs and verbal adjectives from the treebanks for the two languages. As discussed in Section 2, we consider both prepositions in PP modifications as well as preposition complements of verbs. We investigate to what extent the preferences for particular prepositions are shared between a verb and a verbal adjective by using the preposition distribution of the verbal adjective as the reference distribution and the preposition distribution of the verb as a predictor. The particulars of this evaluation will be discussed in more detail in Section 5.

In order to obtain reliable probability distributions from co-occurrence counts, a large number of examples for each verb and verbal adjective is needed. Consequently, this study is conducted using large, machine-annotated treebanks. Such automatic annotations, of course, contain parsing errors, and PP attachment is one of the most frequent attachment errors (Kummerfeld et al., 2012; Mirroshandel et al., 2012; de Kok et al., 2017). However, it should be pointed out that there is far less ambiguity in the attachment of prepositions to verbal adjectives since there is usually no ambiguity in the case of PP modification of prenominal verbal adjective modifiers (see the PP attachment in (2)). For example, the parser of de Kok and Hinrichs (2016) attaches 84.47% of the prepositions that have an attributive adjective as their head correctly. Since verbal adjectives form the reference distribution in our experiments, we are evaluating against a set with fewer attachment errors than the average number of preposition attachment

[2] A more extensive discussion of this type of ambiguity in German can be found in Maienborn (2007). Zwart (2011) provides a more thorough discussion for the phenomenon in Dutch, and we refer to Bresnan (1980) and Levin and Rappaport (1986) for the analysis of adjectival passives in English.

errors. In the remainder of this section, we describe in more detail the Dutch and German data that is used in our study.

Dutch For our study of PP-modification of verbal adjectives in Dutch we use the Lassy Large treebank of written Dutch (Van Noord et al., 2013). Lassy Large consists of approximately 700 million words across various text genres, including newspaper, medical, encyclopedic, and political texts. Each sentence in Lassy Large is syntactically annoted using the Alpino dependency parser (Van Noord, 2006).

The Alpino lexicon encodes adjectives that are derived from past and present participles using lexical tags that indicate their verbal origin. This information percolates to the feature structures and is available in the final XML serialization of the dependency structure. Consequently, verbal adjectives can be extracted using simple attribute-based queries over the Lassy treebank. The extraction is further accommodated by the fact that the Lassy treebank uses the verb infinitive as the lemma for a verbal adjective, as specified by the D-COI annotation guidelines (Van Eynde, 2005) that Lassy uses for tagging and lemmatization. Consequently, there is a one-to-one mapping of verbal adjectives to their corresponding verbs. Since infinitive modifications are considered to be verbs in Alpino, we do not include them in the present study.

We extract verbs and verbal adjectives and the prepositions that they govern with one of the following three dependency relations: (i) prepositional phrase modification (*pp/mod*); (ii) preposition complements (*pp/pc*); and (iii) locative/directional complements (*pp/ld*). For prenominal modifiers, we include modifications using both the categories *ap* and *ppart*. In the extraction, we also consider prepositions that are multi-word units (such as *ten aanzien van* 'with regards to'), multi-headed prepositions, and reentrancies in the dependency structure.

German For our study of PP-modification in German, we extract the relevant data from two sections of the TüBa-D/DP treebank. The first section consists of articles from the German newspaper taz from the period 1986 to 2009 (393.7 million tokens and 28.9 million sentences). The second is based on the German Wikipedia dump of January 1, 2017 (747.7 million tokens and 40.2 million sentences). Both treebanks were annotated using the parser of de Kok and Hinrichs (2016) and then lemmatized using the SepVerb lemmatizer (de Kok, 2014).

In our study, we consider prepositions in (i) prepositional phrase modifications (*PP*) and (ii) prepositional complements *(OBJP)*, along with their respective verb or verbal adjective governor. In contrast to the Dutch treebank where lexical tags indicate an adjective's verbal origin, such information was not available for the German adjectives. In the German treebank, verbal adjectives are lemmatized to their adjective lemmas. For example, *beschrifteter* 'labeled' is lemmatized to *beschriftet* 'labeled'. Therefore, all adjectives are analyzed by the SMOR morphological analyzer (Schmid et al., 2004) in order to detect verbal components in the adjectives. When the SMOR analysis of an adjective reveals components that imply a verbal reading, the forms are labeled as *verb-derived* in the treebank. In addition, the corresponding base verb lemma is reconstructed from the analysis.

In contrast to the Dutch data, the availability of a wide-coverage morphological analyzer has also made it possible to include many adjectives that have transitioned from verbal adjectives to full adjectives in the data set. For instance, the adjective *unbegrenzbar* 'illimitable' is recognized as a verb-derived adjective and lemmatized to the corresponding verb base form *begrenzen* 'to limit'.

Set partitioning As discussed in Section 3, there is an ambiguity between the verbal and adjectival analyses of participles when the participle is governed by a verb form that can both be auxiliary and copular. For this reason, we create three different co-occurrence sets for both Dutch and German: (i) the confusion set of verbs and verbal adjectives that are in such ambiguous positions; (ii) the set of verbs that are not in such ambiguous positions; and (iii) the set of verbal adjectives that are not in such ambiguous positions.

5 Experiments

The goal of our experiments is to test our thesis that there are distributional regularities between verbal adjectives and their corresponding verbs. As motivated in Section 2, we will look at co-occurrences with

prepositions in particular. In our experiments, we will use *relative entropy* (Kullback-Leibler divergence) to determine how much a distribution Q diverges from a reference distribution P (Equation 1).

$$D(P \parallel Q) = \sum_i P(i) \lg \frac{P(i)}{Q(i)} \tag{1}$$

The relative entropy estimates the expected number of additional bits that is required when a sample of P is encoded using a code optimized for Q rather than P. The divergence is zero when the two distributions are identical.

For each subset (Section 4) of our dataset, we estimate a probability distribution $P^*(p|v)$ using maximum likelihood estimation, where p is the preposition, v the verb lemma, and $count(v, p)$ the number of times v governs p with a prepositional phrase or prepositional complement relation in the data set (Equation 2).[3]

$$P^*(p|v) = \frac{\text{count}(v, p)}{\sum_{p^\#} \text{count}(v, p^\#)} \tag{2}$$

The relative entropy for a conditional distribution is the (possibly weighted) average of relative entropies of verbs (Equation 3). However, the average relative entropy obscures the differences in relative entropy between frequent and infrequent lemmas. Instead, we sort verbal lemmas by their frequency in the set from which P derives. We then plot the moving average of maximally 500 lemmas in frequency order.[4] The resulting graph shows the change in relative entropy as the lemmas become more rare.

$$D(P \parallel Q) = \sum_v P(v) \sum_p P(p|v) \lg \frac{P(p|v)}{Q(p|v)} \tag{3}$$

We perform four experiments in total, computing the divergences in Table 1. In each experiment, the verbal adjective set is used as the reference distribution P. This is motivated by the fact that verbal adjectives have fewer PP attachment ambiguities and thus serve as a better reference distribution. Furthermore, since verbs are often far more frequent than verbal adjectives, one would typically want to predict the co-occurrences of a verbal adjective.

Set for P	**Set for** Q
Verbal adjectives (Dutch)	Verbs (Dutch)
Verbal adjectives (German)	Verbs (German)
Ambiguous verbal adjectives/participles (Dutch)	Verbs (Dutch)
Ambiguous verbal adjectives/participles (German)	Verbs (German)

Table 1: The four different pairs of distributions that are evaluated.

We only consider lemmas which occur at least 50 times in each of the paired sets of Table 1. Work on word embeddings has shown that a reasonable number of occurrences is required to get a reliable sample of the contexts in which a word occurs. Consequently, low-frequency words are typically discarded (Collobert et al., 2011; Pennington et al., 2014).

As mentioned before in Section 4 the set of prepositions we consider includes, besides the simplex prepositions in each language, also multi-word units, multi-headed prepositions, etc. The resulting sets of prepositions over which the distributions are computed is relatively large: 1060 prepositions for Dutch and 10,665 prepositions for German. The large proliferation of prepositions has two causes: (i) different spelling variations of prepositions (e.g. *voor* 'for' is sometimes emphasized as *vóór*); and (ii) errors

[3]Note that including verbs that do not govern a preposition in the denominator would result in an improper probability distribution, since then $\sum_p P^*(p|v) \neq 1$. However, the observation made by one reviewer - that they may need to be counted - leads to an interesting question: Do some verbs have a stronger tendency to be modified by prepositional phrases than others, and are these tendencies shared by verbs and their corresponding verb-derived adjectives?

[4]The use of the raw data points results in very uneven graphs.

caused by the automatic annotation. However, since the large majority of prepositions are in the long tail, they have virtually no bearing on the evaluation.[5]

Unconditional model We compare the verb-based distributions with a baseline model that computes unconditional preposition probabilities over a verb set $Q_u(p)$ (Equation 4).

$$Q_u(p) = \frac{\sum_{v^\#} \text{count}(v^\#, p)}{\sum_{v^\#,p^\#} \text{count}(v^\#, p^\#)} \tag{4}$$

Mixture model Since the adjective sets contain deverbal adjectives, we expect the verb models to overestimate the probabilities of prepositions that co-occur with the verbal reading of the adjective. For example, consider the adjective *geschlossen* 'closed' that is discussed in Section 3.2. Because the verb set only contains the verbal reading of *geschlossen*, it will underestimate the probabilities of prepositions that co-occur with the deverbal reading of *geschlossen*. To smoothen the distribution of the verb model, we also introduce a mixture model $Q_m(p|v)$ that combines the verb and unconditional models (Equation 5).

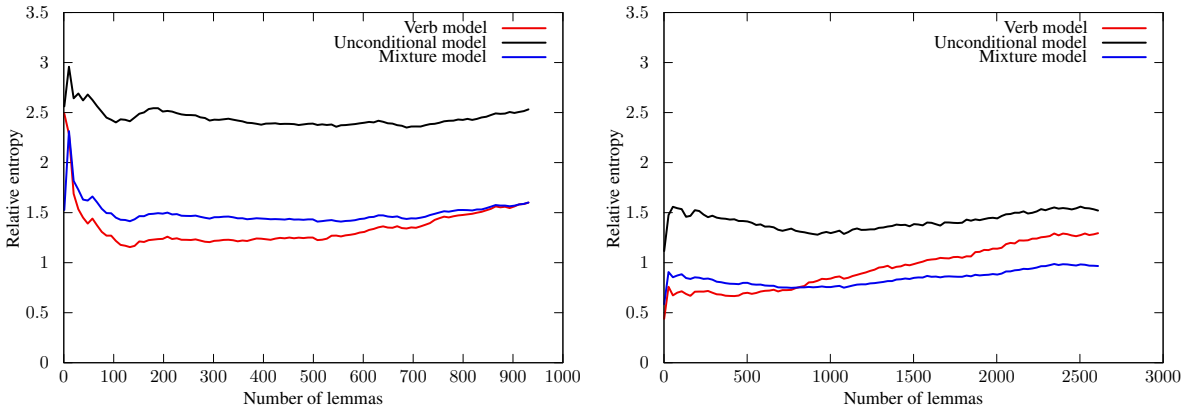
$$Q_m(p|v) = \frac{Q(p|v) + Q_u(p)}{2} \tag{5}$$

In the following section, we report and discuss the results for the experiments described in this section.

6 Main Results and Implications for Computational Modeling

The main result of our experiments is that verbs and verbal adjectives share significant distributional regularities. This permits the distribution of prepositions for verbal adjectives to be reliably predicted using the preposition distributions of their corresponding verbal lemmas. Figure 1a shows, on the Y-axis, the relative entropy of the three different variants of verb-based distributions (introduced in Section 5) and the reference verbal adjective distribution for Dutch. Aside from a small subset of highly frequent verbal adjectives, the verb distribution (red) turns out to be the best predictor of the verbal adjective distribution. For the more infrequent lemmas, however, the performance of the verb model converges towards the performance of the more general mixture model (blue).

Figure 1b presents the same analysis using the German data. The general trend is the same for both Dutch and German: the verb distribution is the best for modeling frequent verbal adjectives (the first 800-900 lemmas).[6] The mixture distribution provides a surprisingly stable approximation, even as the frequency of the verbal adjectives decreases. In both languages the verb and mixture models outperform the unconditional model baseline (black).



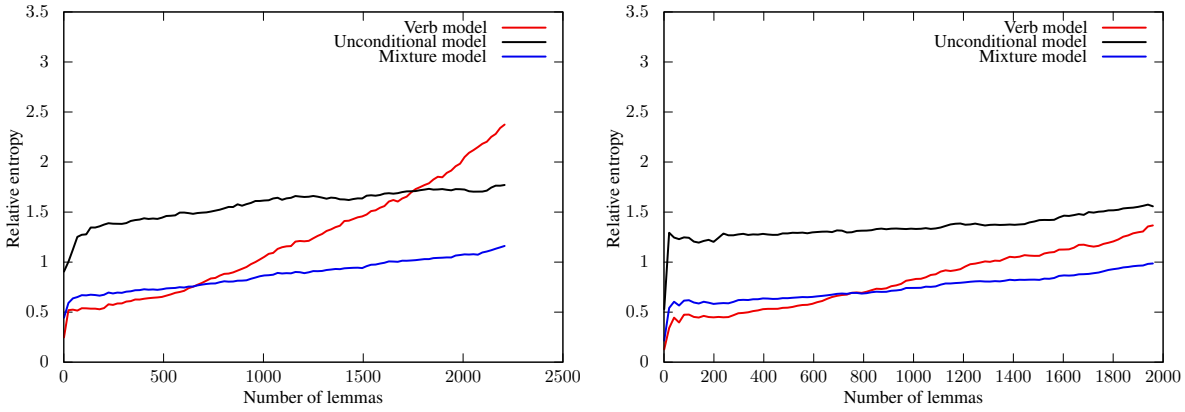
(a) Verbal adjectives (Dutch)

(b) Verbal adjectives (German)

Figure 1: Prediction of prepositions attached to verbal adjectives.

[5]Only 347 of the Dutch prepositions and 690 of the German prepositions occur at least 50 times in our datasets.

[6]The large difference in the number of verbal adjectives in Dutch and German is cause by the fact that for German we also consider verb-derived adjectives like *unbegrenzbar* 'illimitable', see Section 4. These are not considered for Dutch.



(a) Ambiguous verbal adjectives/participles (Dutch)

(b) Ambiguous verbal adjectives/participles (German)

Figure 2: Prediction of prepositions attached to ambiguous verbal adjectives/participles.

Figures 2a and 2b display the relative entropy values obtained by the verb, mixture and unconditional models with respect to the distribution of ambiguous verbal adjectives/participles. The graphs show similar trends as in the unambiguous verbal adjectives case.[7]

To conclude, the case study has shown that there is a significant overlap in the syntagmatic distribution of different morphosyntactic realizations of a verb lemma. To be able to exploit this overlap in distributional and computational modeling, it is crucial that different morphosyntactic realizations of a lexical root are linked to the same lemma. The utility of incorporating sub-word information in distributional modeling has already been recognized and led to the development of character-based representations. However, these representations have been largely constructed on the basis of small supervised training sets. Such small training sets only contain a limited vocabulary, giving representation learners little opportunity to learn the similarities that exist between different morphological realizations of a verbal lemma. As shown by our study, reliable distributions require a reasonably large sample of co-occurrences, which is not provided by such small data sets. The performance of the verb model deteriorates as the number of available samples decreases. In preliminary work, we have seen that a fairly large sample is needed to faithfully model the underlying distribution.

7 Implications for Treebanking

Our study of distributional regularities of verbs and verbal adjectives has shown that treebanks have the potential to contribute to models with good generalization behavior. However, discovering such regularities is greatly helped by providing the necessary annotations in the treebank. In this section, we give a brief overview of which annotations are particularly relevant to the analysis of verbs and verbal adjectives.

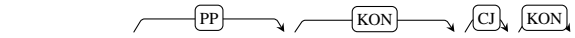
To estimate co-occurrence distributions of verb lemmas and words that enter a dependency relation with them, the verbal and adjectival occurrences of each verb should be annotated with the verbal lemma in a treebank. Even though many treebanks annotate tokens with their lemmas, verbal adjectives are typically lemmatized to their adjectival lemma and not their verbal lemma (see Section 4). In addition, it would be useful if treebanks annotated forms that have fully transitioned into adjectives with their original verb lemma as well.

Another annotation that would have been useful to our study, would be a lexical attribute that indicates whether a verb-derived adjective has a verbal or a deverbal reading. This is particularly useful in cases where verbal and deverbal readings have the same surface form, such as the adjective *geschlossen* that was discussed in Section 3.2. Separation of the verbal readings from the deverbal readings would make it possible to only rely on the verb distribution for predicting the co-occurrences of verbal adjectives.

[7]For the Dutch dataset, the ambiguous verbal adjectives/participles make up 17.05% of the dataset, compared to only 3.49% ambiguous cases for German. The reason is that Dutch has several verbs that have both auxiliary and copular readings, while in German only *sein* 'to-be' can be ambiguous (see Section 3.1).

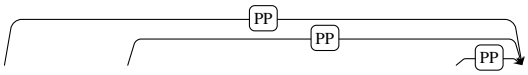
Finally, the extraction of verb and preposition co-occurrences for German was hampered by the annotation of prepositional phrase conjunctions and verb conjunctions. The dependency annotation guidelines (Foth, 2006) use shallow analyses of conjunctions, including PP conjunctions, such the one in (5a). A deeper structure needs to be constructed to infer that the second occurrence of the preposition *über* 'about' is also governed by *ärgert* 'agitates'. Conversely, a prepositional phrase can be governed by more than one verb or verbal adjective, as shown in (5b). However, such annotations are not possible in the German treebank that we used, since the annotation guidelines adhere to the single-headedness principle. Deeper annotations, such as those provided in the Lassy Large treebank - which was automatically annotated using the Alpino parser for Dutch (Van Noord, 2006) - help tremendously in exhaustive co-occurrence extraction.

(5) a.



Staffelt ärgert sich über den Lärm und auch über Senator Haase
Staffelt agitates himself about the noise and also about senator Haase

b.



Vertaald , ingeleid en van toelichtingen voorzien door H. Savenije
Translated , prefaced and of comments supplied by H. Savenije

Acknowledgments

Financial support for the research reported in this paper was provided by the German Research Foundation (DFG) as part of the Collaborative Research Center "The Construction of Meaning" (SFB 833), project A3.

References

Joan Wanda Bresnan. 1980. *The passive in lexical theory*. Massachusetts Institute of Technology, Center for Cognitive Science.

Noam Chomsky. 1970. Remarks on Nominalization. In Roderick A. Jacobs and Peter S. Rosenbaum, editors, *Readings in English Transformational Grammar*. Ginn, Boston, pages 184–221.

Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. 2011. Natural Language Processing (almost) From Scratch. *Journal of Machine Learning Research* 12(Aug):2493–2537.

Daniël de Kok. 2014. TüBa-D/W: a large dependency treebank for German. In *Proceedings of the 13th International Workshop on Treebanks and Linguistic Theories*. Tübingen, Germany, pages 271–278.

Daniël de Kok and Erhard Hinrichs. 2016. Transition-based dependency parsing with topological fields. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Association for Computational Linguistics, Berlin, Germany, pages 1–7. http://anthology.aclweb.org/P16-2001.

Daniël de Kok, Jianqiang Ma, Corina Dima, and Erhard Hinrichs. 2017. PP Attachment: Where do We Stand? In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL): Volume 2, Short Papers*. Valencia, Spain, pages 311–317. http://www.aclweb.org/anthology/E17-2050.

John Rupert Firth. 1957. A synopsis of linguistic theory 1930-1955. In *Studies in Linguistic Analysis (special volume of the Philological Society)*. The Philological Society, Oxford, pages 1–32.

Kilian A. Foth. 2006. Eine Umfassende Constraint-Dependenz-Grammatik Des Deutschen .

Dale Gerdemann. 1994. Complement Inheritance as Subcategorization Inheritance. In John Nerbonne, Carl Pollard, and Klaus Netter, editors, *German in Head-Driven Phrase Structure Grammar*. Center for Study of Language and Information: Stanford University, CSLI Lecture Notes, pages 341–363.

Zelig Harris. 1951. *Methods in Structural Linguistics*. University of Chicago Press: Chicago.

Jonathan K. Kummerfeld, David Hall, James R. Curran, and Dan Klein. 2012. Parser showdown at the Wall Street corral: An empirical investigation of error types in parser output. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Association for Computational Linguistics, Jeju Island, Korea, pages 1048–1059. http://www.aclweb.org/anthology/D12-1096.

Robert B. Lees. 1960. The Grammar of English Nominalizations. *International Journal of American Linguistics* 26(3). Part II.

Barbara Lenz. 1993. Probleme Der Kategorisierung Deutscher Partizipien. *Zeitschrift für Sprachwissenschaft* 12(1):39–76.

Beth Levin and Malka Rappaport. 1986. The Formation of Adjectival Passives. *Linguistic inquiry* 17(4):623–661.

Claudia Maienborn. 2007. Das Zustandspassiv. Grammatische Einordnung–Bildungsbeschränkung–Interpretationsspielraum. *Zeitschrift für germanistische Linguistik* 35(1-2):83–114.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems (NIPS)*. Lake Tahoe, Nevada, United States, pages 3111–3119.

Seyed Abolghasem Mirroshandel, Alexis Nasr, and Joseph Le Roux. 2012. Semi-supervised Dependency Parsing using Lexical Affinities. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Jeju Island, Korea, pages 777–785. http://www.aclweb.org/anthology/P12-1082.

Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. GloVe: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar, pages 1532–1543.

Helmut Schmid, Arne Fitschen, and Ulrich Heid. 2004. SMOR: A German computational morphology covering derivation, composition, and inflection. In *Proceedings of the IVth International Conference on Language Resources and Evaluation (LREC 2004)*. Lisbon, Portugal, pages 1263–1266.

Frank Van Eynde. 2005. Part of Speech Tagging En Lemmatisering Van Het D-Coi Corpus. *Intermediate, project internal version* .

Gertjan Van Noord. 2006. At last parsing is now operational. In *TALN06. Verbum Ex Machina. Actes de la 13e conference sur le traitement automatique des langues naturelles*. Leuven, Belgium, pages 20–42.

Gertjan Van Noord, Gosse Bouma, Frank Van Eynde, Daniël De Kok, Jelmer Van der Linde, Ineke Schuurman, Erik Tjong Kim Sang, and Vincent Vandeghinste. 2013. Large scale syntactic annotation of written Dutch: Lassy. In *Essential Speech and Language Technology for Dutch*, Springer, pages 147–164.

Jan-Wouter Zwart. 2011. *The syntax of Dutch*. Cambridge Syntax Guides. Cambridge University Press.