

A Dataset for Joint Noun–Noun Compound Bracketing and Interpretation

Murhaf Fares

Department of Informatics

University of Oslo

murhaff@ifi.uio.no

Abstract

We present a new, sizeable dataset of noun–noun compounds with their syntactic analysis (bracketing) and semantic relations. Derived from several established linguistic resources, such as the Penn Treebank, our dataset enables experimenting with new approaches towards a holistic analysis of noun–noun compounds, such as joint-learning of noun–noun compounds bracketing and interpretation, as well as integrating compound analysis with other tasks such as syntactic parsing.

1 Introduction

Noun–noun compounds are abundant in many languages, and English is no exception. According to Ó Séaghdha (2008), three percent of all words in the British National Corpus (Burnard, 2000, BNC) are part of nominal compounds. Therefore, in addition to being an interesting linguistic phenomenon per se, the analysis of noun–noun compounds is important to other natural language processing (NLP) tasks such as machine translation and information extraction. Indeed, there is already a nontrivial amount of research on noun–noun compounds within the field of computational linguistics (Lauer, 1995; Nakov, 2007; Ó Séaghdha, 2008; Tratz, 2011, inter alios).

As Lauer and Dras (1994) point out, the treatment of noun–noun compounds involves three tasks: identification, bracketing and semantic interpretation. With a few exceptions (Girju et al., 2005; Kim and Baldwin, 2013), most studies on noun–noun compounds focus on one of the aforementioned tasks in isolation, but these tasks are of course not fully independent and therefore might benefit from a joint-learning approach, especially bracketing and semantic interpretation.

Reflecting previous lines of research, most of the existing datasets on noun–noun compounds either include bracketing information or semantic relations, rarely both. In this article we present a fairly large dataset for noun–noun compound bracketing as well as semantic interpretation. Furthermore, most of the available datasets list the compounds out of context. Hence they implicitly assume that the semantics of noun–noun compounds is type-based; meaning that the same compound will always have the same semantic relation. To test this assumption of type-based vs. token-based semantic relations, we incorporate the context of the compounds in our dataset and treat compounds as tokens rather than types. Lastly, to study the effect of noun–noun compound bracketing and interpretation on other NLP tasks, we derive our dataset from well-established resources that annotate noun–noun compounds as part of other linguistic structures, viz. the Wall Street Journal Section of the Penn Treebank (Marcus et al., 1993, PTB), PTB noun phrase annotation by Vadas and Curran (2007), DeepBank (Flickinger et al., 2012), the Prague Czech–English Dependency Treebank 2.0 (Hajič et al., 2012, PCEDT) and NomBank (Meyers et al., 2004). We therefore can quantify the effect of compound bracketing on syntactic parsing using the PTB, for example.

In the following section, we review some of the existing noun compound datasets. In § 3, we present the process of constructing a dataset of noun–noun compounds with bracketing information and semantic relations. In § 4, we explain how we construct the bracketing of noun–noun compounds from three resources and report ‘inter-resource’ agreement levels. In § 5, we present the semantic relations extracted from two resources and the correlation between the two sets of relations. In § 6, we conclude the article and present an outlook for future work.

Dataset	Size	Relations	Bracketing
Nastase & Szpakowicz	600	30	No
Girju et al.	4,500	21	600
Ó Séaghdha & Copestake	1,443	6	No
Kim & Baldwin ¹	2,169	20	No
Tratz & Hovy	17,509	43	No

Table 1: Overview of noun compound datasets. Size: type count

2 Background

The syntax and semantics of noun–noun compounds have been under focus for years, in linguistics and computational linguistics. Levi (1978) presents one of the early and influential studies on noun–noun compounds as a subset of so-called complex nominals. Levi (1978) defines a set of nine “recoverably deletable predicates” which express the “semantic relationship between head nouns and prenominal modifiers” in complex nominals. Finin (1980) presented one of the earliest studies on nominal compounds in computational linguistics, but Lauer (1995) was among the first to study statistical methods for noun compound analysis. Lauer (1995) used the Grolier encyclopedia to estimate word probabilities, and tested his models on a dataset of 244 three-word bracketed compounds and 282 two-word compounds. The compounds were annotated with eight prepositions which Lauer takes to approximate the semantics of noun–noun compounds.

Table 1 shows an overview of some of the existing datasets for nominal compounds. The datasets by Nastase and Szpakowicz (2003) and Girju et al. (2005) are not limited to noun–noun compounds; the former includes compounds with adjectival and adverbial modifiers, and the latter has many noun-preposition-noun constructions. The semantic relations in Ó Séaghdha and Copestake (2007) and Kim and Baldwin (2008) are based on the relations introduced by Levi (1978) and Barker and Szpakowicz (1998), respectively. All of the datasets in Table 1 list the compounds out of context. In addition, the dataset by Girju et al. (2005) includes three-word bracketed compounds, whereas the rest include two-word compounds only. On the other hand, (Girju et al., 2005) is the only dataset in Table 1 that is not publicly available.

¹In Table 1 we refer to (Kim and Baldwin, 2008), the other dataset by Kim and Baldwin (2013), which includes 1,571 three-word compounds, is not publicly available.

	NNP ^h	NNP ⁰
Compounds	38,917	29,666
Compound types	21,016	14,632

Table 2: Noun–noun compounds in WSJ Corpus

3 Framework

This section gives an overview of our method to automatically construct a bracketed and semantically annotated dataset of noun–noun compounds from four different linguistic resources. The construction method consists of three steps that correspond to the tasks defined by Lauer and Dras (1994): identification, bracketing and semantic interpretation.

Firstly, we identify the noun–noun compounds in the PTB WSJ Section using two of the compound identification heuristics introduced by Fares et al. (2015), namely the so-called syntax-based NNP^h heuristic which includes compounds that contain common and proper nouns but excludes the ones headed by proper nouns, and the syntax-based NNP⁰ heuristic which excludes all compounds that contain proper nouns, be it in the head position or the modifier position. Table 2 shows the number of compounds and compound types we identified using the NNP^h and NNP⁰ heuristics. Note that the number of compounds will vary in the following sections depending on the resources we use.

Secondly, we extract the bracketing of the identified compounds from three resources: PTB noun phrase annotation by Vadas and Curran (2007), DeepBank and PCEDT. Vadas and Curran (2007) manually annotated the internal structure of noun phrases (NPs) in PTB which were originally left unannotated. However, as is the case with other resources, Vadas and Curran (2007) annotation is not completely error-free, as shown by Fares et al. (2015). We therefore crosscheck their bracketing through comparing to those of DeepBank and PCEDT. The latter two, however, do not contain explicit annotation of noun–noun compound bracketing, but we can ‘reconstruct’ the bracketing based on the dependency relations assigned in both resources, i.e. the logical form meaning representation in DeepBank and the tectogrammatical layer (*t*-layer) in PCEDT. Based on the bracketing extracted from the three resources, we define the subset of compounds that are bracketed similarly in the three resources. Lastly, we extract the se-

semantic relations of two-word compounds as well as multi-word bracketed compounds from two resources: PCEDT and NomBank.

On a more technical level, we use the so-called phrase-structure layer (p -layer) in PCEDT to identify noun–noun compounds, because it includes the NP annotation by Vadas and Curran (2007), which is required to apply the noun–noun compound identification heuristics by Fares et al. (2015). For bracketing, we also use the PCEDT p -layer, in addition to the dataset prepared by Oepen et al. (2016) which includes DeepBank and the PCEDT tectogrammatical layer. We opted for the dataset by Oepen et al. (2016) because they converted the tectogrammatical annotation in PCEDT to dependency representation in which the “set of graph nodes is equivalent to the set of surface tokens.” For semantic relations, we also use the dataset by Oepen et al. (2016) for PCEDT relations and the original NomBank files for NomBank relations.

Throughout the whole process we store the data in a relational database with a schema that represents the different types of information, and the different resources from which they are derived. As we will show in § 4 and § 5, this set-up allows us to combine information in different ways and therefore create ‘different’ datasets.

4 Bracketing

Noun–noun compound bracketing can be defined as the disambiguation of the internal structure of compounds with three nouns or more. For example, we can bracket the compound *noon fashion show* in two ways:

1. Left-bracketing: [[noon fashion] show]
2. Right-bracketing: [noon [fashion show]]

In this example, the right-bracketing interpretation (*a fashion show happening at noon*) is more likely than the left-bracketing one (*a show of noon fashion*). However, the correct bracketing need not always be as obvious, some compounds can be subtler to bracket, e.g. *car radio equipment* (Girju et al., 2005).

4.1 Data & Results

As explained in § 3, we first identify noun–noun compounds in the WSJ Corpus, then we extract and map their bracketing from three linguistic resources: PCEDT, DeepBank and noun phrase annotation by Vadas and Curran (2007) (VC-PTB,

henceforth). Even though we can identify 38,917 noun–noun compounds in the full WSJ Corpus (cf. Table 2), the set of compounds that constitutes the basis for bracketing analysis (i.e. the set of compounds that occur in the three resources) is smaller. First, because DeepBank only annotates the first 22 Sections of the WSJ Corpus. Second, because not all the noun sequences identified as compounds in VC-PTB are treated as such in DeepBank and PCEDT. Hence, the number of compounds that occur in the three resources is 26,500. Furthermore, three-quarters (76%) of these compounds consist of two nouns only, meaning that they do not require bracketing, which leaves us a subset of 6,244 multi-word compounds—we will refer to this subset as the bracketing subset.

After mapping the bracketings from the three resources we find that they agree on the bracketing of almost 75% of the compounds in the bracketing subset. Such an agreement level is relatively good compared to previously reported agreement levels on much smaller datasets, e.g. Girju et al. (2005) report a bracketing agreement of 87% on a set of 362 three-word compounds. Inspecting the disagreement among the three resources reveals two things. First, noun–noun compounds which contain proper nouns (NNP) constitute 45% of the compounds that are bracketed differently. Second, 41% of the differently bracketed compounds are actually sub-compounds of larger compounds. For example, the compound *consumer food prices* is left-bracketed in VC-PTB, i.e. [[*consumer food*] *prices*], whereas in PCEDT and DeepBank it is right-bracketed. This difference in bracketing leads to two different sub-compounds, namely *consumer food* in VC-PTB and *food prices* in PCEDT and DeepBank.

It is noteworthy that those two observations do not reflect the properties of compounds containing proper nouns or sub-compounds; they only tell us their percentages in the set of differently bracketed compounds. In order to study their properties, we need to look at the number of sub-compounds and compounds containing NNPs in the set of compounds where the three resources agree. As it turns out, 72% of the compounds containing proper nouns and 76% of the sub-compounds are bracketed similarly. Therefore when we exclude them from the bracketing subset we do not see a significant change in bracketing agreement among

	$\alpha-\beta$	$\alpha-\gamma$	$\beta-\gamma$	$\alpha-\beta-\gamma$
NNP ^h	80%	79%	88%	75%
NNP ⁰	78%	75%	90%	74%
NNP ^h w/o sub	82%	82%	86%	75%
NNP ⁰ w/o sub	81%	77%	90%	74%

Table 3: Bracketing agreement – α : DeepBank; β : PCEDT; γ : VC-PTB; NNP⁰: excl. proper nouns; NNP^h: incl. proper nouns; w/o sub: excl. sub-compounds

the three resources, as shown in the right-most column in Table 3.

We report pairwise bracketing agreement among the three resources in Table 3. We observe higher agreement level between PCEDT and VC-PTB than the other two pairs; we speculate that the annotation of the t -layer in PCEDT might have been influenced by the so-called phrase-structure layer (p -layer) which in turn uses VC-PTB annotation. Further, PCEDT and VC-PTB seem to disagree more on the bracketing of noun–noun compounds containing NNPs; because when proper nouns are excluded (NNP⁰), the agreement level between PCEDT and VC-PTB increases, but it decreases for the other two pairs.

As we look closer at the compound instances where at least two of the three resources disagree, we find that some instances are easy to classify as annotation errors. For example, the compound *New York streets* is bracketed as right-branching in VC-PTB, but we can confidently say that this is a left-bracketing compound. Not all bracketing disagreements are that easy to resolve though; one example where left- and right-bracketing can be accepted is *European Common Market approach*, which is bracketed as follows in DeepBank (1) and PCEDT and VC-PTB (2):

1. [[European [Common Market]] approach]
2. [European [[Common Market] approach]]

Even though this work does not aim to resolve or correct the bracketing disagreement between the three resources, we will publish a tool that allows resource creators to inspect the bracketing disagreement and possibly correct it.

5 Relations

Now that we have defined the set of compounds whose bracketing is agreed-upon in different resources, we move to adding semantic relations to

Compound	Functor	NomBank Arg
Negligence penalty	CAUS	ARG3
Death penalty	RSTR	ARG2
Staff lawyer	RSTR	ARG3
Government lawyer	APP	ARG2

Table 4: Example compounds with semantic relations

our dataset. We rely on PCEDT and NomBank to define the semantic relations in our dataset, which includes bracketed compounds from § 4 as well as two-word compounds. However, unlike § 4, our set of noun–noun compounds in this section consists of the compounds that are bracketed similarly in PCEDT and VC-PTB and occur in both resources.² This set consists of 26,709 compounds and 14,405 types.

PCEDT assigns syntactico-semantic labels, so-called functors, to all the syntactic dependency relations in the tectogrammatical layer (a deep syntactic structure). Drawing on the valency theory of the Functional Generative Description, PCEDT defines 69 functors for verbs as well as nouns and adjectives (Cinková et al., 2006).³ NomBank, on the other hand, is about nouns only; it assigns role labels (arguments) to common nouns in the PTB. In general, NomBank distinguishes between predicate arguments and modifiers (adjuncts) which correspond to those defined in PropBank (Kingsbury and Palmer, 2002).⁴ We take both types of roles to be part of the semantic relations of noun–noun compounds in our dataset.

Table 4 shows some examples of noun–noun compounds annotated with PCEDT functors and NomBank arguments. The functor CAUS expresses causal relationship; RSTR is an underspecified adnominal functor that is used whenever the semantic requirements for other functors are not met; APP expresses appurtenance. While the PCEDT functors have specific definitions, most of the NomBank arguments have to be interpreted in connection with their predicate or frame. For ex-

²We do not use the intersection of the three resources as in § 4, because DeepBank does not contribute to the semantic relations of noun–noun compounds and it limits the size of our dataset (cf. § 4). Nonetheless, given our technical set-up we can readily produce the set of compounds that occur in the three resources and are bracketed similarly, and then extract their semantic relations from PCEDT and NomBank.

³The full inventory of functors is available on <https://ufal.mff.cuni.cz/pcedt2.0/en/functors.html> (visited on 22/04/2016).

⁴See Table 2 in Meyers (2007, p. 90) for an overview of adjunct roles in NomBank.

ample, ARG3 of the predicate *penalty* in Table 4 describes crime whereas ARG3 of the predicate *lawyer* describes rank. Similarly, ARG2 in *penalty* describes punishment, whereas ARG2 in *lawyer* describes beneficiary or consultant.

5.1 Data & Results

Given 26,709 noun–noun compounds, we construct a dataset with two relations per compound: a PCEDT functor and a NomBank argument. The resulting dataset is relatively large compared to the datasets in Table 1. However, the largest dataset in Table 1, by Tratz and Hovy (2010), is type-based and does not include proper nouns. The size of our dataset becomes 10,596 if we exclude the compounds containing proper nouns and only count the types in our dataset; this is still a relatively large dataset and it has the important advantage of including bracketing information of multi-word compounds, *inter alia*.

Overall, the compounds in our dataset are annotated with 35 functors and 20 NomBank arguments, but only twelve functors and nine NomBank arguments occur more than 100 times in the dataset. Further, the most frequent NomBank argument (ARG1) accounts for 60% of the data, and the five most frequent arguments account for 95%. We see a similar pattern in the distribution of PCEDT functors, where 49% of the compounds are annotated with RSTR (the least specific adnominal functor in PCEDT). Further, the five most frequent functors account for 89% of the data (cf. Table 5). Such distribution of relations is not unexpected because according to Cinková et al. (2006), the relations that cannot be expressed by “semantically expressive” functors usually receive the functor PAT, which is the second most frequent functor. Furthermore, Kim and Baldwin (2008) report that 42% of the compounds in their dataset are annotated as TOPIC, which appears closely related to ARG1 in NomBank.

In theory, some of the PCEDT functors and NomBank arguments express the same type of relations. We therefore show the ‘correlation’ between PCEDT functors and NomBank arguments in Table 5. The first half of the table maps PCEDT functors to NomBank arguments, and the second half shows the mapping from NomBank to PCEDT. Due to space limitations, the table only includes a subset of the relations—the most frequent ones. The underlined num-

bers in Table 5 indicate the functors and NomBank arguments that are semantically comparable; for example, the temporal and locative functors (TWHEN, THL, TFRWH and LOC) intuitively correspond to the temporal and locative modifiers in NomBank (ARGM-TMP and ARGM-LOC), and this correspondence is also evident in the figures in Table 5. The same applies to the functor AUTH (authorship) which always maps to the NomBank argument ARG0 (agent). However, not all ‘theoretical similarities’ are necessarily reflected in practice, e.g. AIM vs. ARGM-PNC in Table 5 (both express purpose). NomBank and PCEDT are two different resources that were created with different annotation guidelines and by different annotators, and therefore we cannot expect perfect correspondence between PCEDT functors and NomBank arguments.

PCEDT often assigns more than one functor to different instances of the same compound. In fact, around 13% of the compound types were annotated with more than one functor in PCEDT, whereas only 1.3% of our compound types are annotated with more than one argument in NomBank. For example, the compound *takeover bid*, which occurs 28 times in our dataset, is annotated with four different functors in PCEDT, including AIM and RSTR, whereas in NomBank it is always annotated as ARGM-PNC. This raises the question whether the semantics of noun–noun compounds varies depending on their context, i.e. token-based vs. type-based relations. Unfortunately we cannot answer this question based on the variation in PCEDT because its documentation clearly states that “[t]he annotators tried to interpret complex noun phrases with semantically expressive functors as much as they could. This annotation is, of course, very inconsistent.”⁵ Nonetheless, our dataset still opens the door to experimenting with learning PCEDT functors, and eventually determining whether the varied functors are mere inconsistencies or there is more to this than meets the eye.

6 Conclusion & Future Work

In this article we presented a new noun–noun compound dataset constructed from different linguistic resources, which includes bracketing information and semantic relations. In § 4, we explained

⁵<https://ufal.mff.cuni.cz/pcedt2.0/en/valency.html> (visited on 22/04/2016).

P \ N									Count	Freq
	ARG1	ARG2	ARG0	ARG3	M-LOC	M-MNR	M-TMP	M-PNC		
RSTR	0.60	0.12	0.08	0.10	0.03	0.03	0.01	0.01	12992	48.64
PAT	<u>0.89</u>	0.05	0.01	0.03	0.01	0.01		0.01	3867	14.48
APP	0.42	0.37	0.17	0.01	0.03	0.00	0.00	0.00	3543	13.27
REG	0.75	0.09	0.07	0.07	0.00	0.01	0.00	0.00	2176	8.15
ACT	0.46	0.03	<u>0.48</u>	0.01	0.01	0.00			1286	4.81
LOC	0.16	0.20	<u>0.09</u>	0.01	<u>0.54</u>				979	3.67
TWHEN	0.12	0.04	0.00		<u>0.01</u>		<u>0.81</u>		367	1.37
AIM	0.65	0.12	0.06	0.08	0.00	0.00		<u>0.05</u>	284	1.06
ID	0.39	0.30	0.27	0.04	0.00				256	0.96
MAT	0.86	0.09	0.01	0.02					136	0.51
NE	0.32	0.46	0.13	0.02	0.06				132	0.49
ORIG	0.20	0.19	0.13	0.37	0.06	0.01		0.01	114	0.43
MANN	0.23	0.07	0.01		0.04	<u>0.65</u>			83	0.31
MEANS	0.45	0.09	0.04	0.12	0.14	0.11			56	0.21
EFF	0.60	0.18	0.11	0.04				0.04	55	0.21
AUTH			<u>1.00</u>						49	0.18
BEN	0.45	0.35	0.03	0.17					40	0.15
THL		0.03		0.03			<u>0.95</u>		38	0.14
	ARG1	ARG2	ARG0	ARG3	M-LOC	M-MNR	M-TMP	M-PNC		
RSTR	0.50	0.40	0.38	0.76	0.37	0.79	0.27	0.66		
PAT	<u>0.22</u>	0.05	0.02	0.06	0.02	0.07		0.13		
APP	<u>0.09</u>	0.34	0.22	0.02	0.09	0.00	0.01	0.01		
REG	0.10	0.05	0.05	0.08	0.01	0.02	0.01	0.07		
ACT	0.04	0.01	<u>0.23</u>	0.00	0.02	0.01				
LOC	0.01	0.05	0.03	0.00	<u>0.47</u>					
TWHEN	0.00	0.00	0.00		0.00		<u>0.58</u>			
AIM	0.01	0.01	0.01	0.01	0.00	0.00		<u>0.09</u>		
ID	0.01	0.02	0.03	0.01	0.00					
MAT	0.01	0.00	0.00	0.00						
NE	0.00	0.02	0.01	0.00	0.01					
ORIG	0.00	0.01	0.01	0.02	0.01	0.00		0.01		
MANN	0.00	0.00	0.00		0.00	<u>0.10</u>				
MEANS	0.00	0.00	0.00	0.00	0.01	0.01				
EFF	0.00	0.00	0.00	0.00				0.01		
AUTH			<u>0.02</u>							
BEN	0.00	0.00	<u>0.00</u>	0.00						
THL		0.00		0.00			<u>0.07</u>			
Count	15811	3779	2701	1767	1131	563	510	149		
Freq	59.20	14.15	10.11	6.62	4.23	2.11	1.91	0.56		

Table 5: Correlation between NomBank arguments and PCEDT functors

the construction of a set of bracketed multi-word noun–noun compounds from the PTB WSJ Corpus, based on the NP annotation by Vadas and Curran (2007), DeepBank and PCEDT. In § 5, we constructed a variant of the set in § 4 whereby each compound is assigned two semantic relations, a PCEDT functor and NomBank argument. Our dataset is the largest data set that includes both compound bracketing and semantic relations, and the second largest dataset in terms of the number of compound types excluding compounds that contain proper nouns.

Our dataset has been derived from different resources that are licensed by the Linguistic Data Consortium (LDC). Therefore, we are investigating the possibility of making our dataset publicly available in consultation with the LDC. Otherwise the dataset will be published through the LDC.

In follow-up work, we will enrich our dataset by mapping the compounds in our dataset to the datasets by Kim and Baldwin (2008) and Tratz and Hovy (2010); all of the compounds in the former and some of the compounds in the latter are extracted from the WSJ Corpus. Further, we will experiment with different classification and ranking approaches to bracketing and semantic interpretation of noun–noun compounds using different combinations of relations. We will also study the use of machine learning models to jointly bracket and interpret noun–noun compounds. Finally, we aim to study noun–noun compound identification, bracketing and interpretation in an integrated setup, by using syntactic parsers to solve the identification and bracketing tasks, and semantic parsers to solve the interpretation task.

Acknowledgments. The author wishes to thank Stephan Oepen and Erik Velldal for their helpful assistance and guidance, as well as Michael Roth and the three anonymous reviewers for thoughtful comments. The creation of the new dataset wouldn't have been possible without the efforts of the resource creators from which the dataset was derived.

References

- Ken Barker and Stan Szpakowicz. 1998. Semi-Automatic Recognition of Noun Modifier Relationships. In *Proceedings of the 17th International Conference on Computational Linguistics and the 36th Meeting of the Association for Computational Linguistics*, page 96–102, Montreal, Quebec, Canada.
- Lou Burnard. 2000. Reference guide for the British National Corpus version 1.0.
- Silvie Cinková, Jan Hajič, Marie Mikulová, Lucie Mladová, Anja Nedolužko, Petr Pajas, Jarmila Panevová, Jiří Semecký, Jana Šindlerová, Josef Toman, Zdeňka Urešová, and Zdeněk Žabokrtský. 2006. Annotation of English on the tectogrammatical level: reference book. Technical report, Charles University, Prague. version 1.0.1.
- Murhaf Fares, Stephan Oepen, and Erik Velldal. 2015. Identifying Compounds: On The Role of Syntax. In *International Workshop on Treebanks and Linguistic Theories*, page 273–283, Warsaw, Poland.
- Timothy Wilking Finin. 1980. *The Semantic Interpretation of Compound Nominals*. PhD thesis, University of Illinois at Urbana-Champaign.
- Dan Flickinger, Yi Zhang, and Valia Kordoni. 2012. DeepBank. A dynamically annotated treebank of the Wall Street Journal. In *Proceedings of the 11th International Workshop on Treebanks and Linguistic Theories*, page 85–96, Lisbon, Portugal. Edições Colibri.
- Roxana Girju, Dan Moldovan, Marta Tatu, and Daniel Antohe. 2005. On the semantics of noun compounds. *Computer Speech & Language*, 19(4):479–496.
- Jan Hajič, Eva Hajičová, Jarmila Panevová, Petr Sgall, Ondřej Bojar, Silvie Cinková, Eva Fučková, Marie Mikulová, Petr Pajas, Jan Popelka, Jiří Semecký, Jana Šindlerová, Jan Štěpánek, Josef Toman, Zdeňka Urešová, and Zdeněk Žabokrtský. 2012. Announcing Prague Czech-English Dependency Treebank 2.0. In *Proceedings of the 8th International Conference on Language Resources and Evaluation*, page 3153–3160, Istanbul, Turkey.
- Su Nam Kim and Timothy Baldwin. 2008. Standardised Evaluation of English Noun Compound Interpretation. In *Proceedings of the LREC Workshop: Towards a Shared Task for Multiword Expressions*, page 39–42, Marrakech, Morocco.
- Su Nam Kim and Timothy Baldwin. 2013. A lexical semantic approach to interpreting and bracketing English noun compounds. *Natural Language Engineering*, 19(03):385–407.
- Paul Kingsbury and Martha Palmer. 2002. From TreeBank to PropBank. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, page 1989–1993, Las Palmas, Spain.
- Mark Lauer and Mark Dras. 1994. A probabilistic model of compound nouns. In *Proceedings of the 7th Australian Joint Conference on AI*, page 474–481, Armidale, Australia.
- Mark Lauer. 1995. *Designing Statistical Language Learners. Experiments on Noun Compounds*. Doctoral dissertation, Macquarie University, Sydney, Australia.
- Judith N Levi. 1978. *The syntax and semantics of complex nominals*. Academic Press.
- Mitchell Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. Building a large annotated corpora of English. The Penn Treebank. *Computational Linguistics*, 19:313–330.
- Adam Meyers, Ruth Reeves, Catherine Macleod, Rachel Szekely, Veronika Zielinska, Brian Young, and Ralph Grishman. 2004. Annotating noun argument structure for NomBank. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, page 803–806, Lisbon, Portugal.
- Adam Meyers. 2007. Annotation guidelines for NomBank-noun argument structure for PropBank. Technical report, New York University.
- Preslav Ivanov Nakov. 2007. *Using the Web as an Implicit Training Set: Application to Noun Compound Syntax and Semantics*. Doctoral dissertation, EECS Department, University of California, Berkeley.
- Vivi Nastase and Stan Szpakowicz. 2003. Exploring Noun-Modifier Semantic Relations. In *Fifth International Workshop on Computational Semantics*, page 285–301.
- Diarmuid Ó Séaghdha and Ann Copestake. 2007. Co-occurrence Contexts for Noun Compound Interpretation. In *Proceedings of the Workshop on A Broader Perspective on Multiword Expressions*, page 57–64, Prague, Czech Republic. Association for Computational Linguistics.
- Diarmuid Ó Séaghdha. 2008. Learning compound noun semantics. Technical Report UCAM-CL-TR-735, University of Cambridge, Computer Laboratory, Cambridge, UK.

Stephan Oepen, Marco Kuhlmann, Yusuke Miyao, Daniel Zeman, Silvie Cinková, Dan Flickinger, Jan Hajič, Angelina Ivanova, and Zdeňka Urešová. 2016. Towards Comparability of Linguistic Graph Banks for Semantic Parsing. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC)*, page 3991–3995, Portorož, Slovenia. European Language Resources Association.

Stephen Tratz and Eduard Hovy. 2010. A taxonomy, dataset, and classifier for automatic noun compound interpretation. In *Proceedings of the 48th Meeting of the Association for Computational Linguistics*, page 678–687, Uppsala, Sweden.

Stephen Tratz. 2011. *Semantically-enriched parsing for natural language understanding*. Doctoral dissertation, University of Southern California.

David Vadas and James Curran. 2007. Adding Noun Phrase Structure to the Penn Treebank. In *Proceedings of the 45th Meeting of the Association for Computational Linguistics*, page 240–247, Prague, Czech Republic.