Proceedings of the Australasian Language Technology Workshop 2007, pages 49-56

Extending Sense Collocations in Interpreting Noun Compounds

Su Nam Kim,[♠♡] **Meladel Mistica**[♠] and **Timothy Baldwin**[♠♡]
♠ CSSE
and
♡ NICTA VRL
University of Melbourne
VIC 3010, Australia
{snkim,mmistica,tim}@csse.unimelb.edu.au

Abstract

This paper investigates the task of noun compound interpretation, building on the sense collocation approach proposed by Moldovan et al. (2004). Our primary task is to evaluate the impact of similar words on the sense collocation method, and decrease the sensitivity of the classifiers by expanding the range of sense collocations via different semantic relations. Our method combines hypernyms, hyponyms and sister words of the component nouns, based on WORDNET. The data used in our experiments was taken from the nominal pair interpretation task of SEMEVAL-2007 (4th International Workshop on Semantic Evaluation 2007). In our evaluation, we test 7-way and 2-way class data, and show that the inclusion of hypernyms improves the performance of the sense collocation method, while the inclusion of hyponym and sister word information leads to a deterioration in performance.

1 Introduction

This paper investigates the automatic interpretation of noun compounds (i.e. NCs), such as *paper submission* and *computer science department*. NC interpretation is a well-known problem that aims to predict the **semantic relation** (i.e. SR) between a head noun and its modifying nominal(s) (i.e. modifiers). SRs, simply put, encapsulate how the head noun and the other nominals in a noun compound are related. As English noun phrases are right-headed, the head noun occurs after all modifying nouns. For example, *brick house* is interpreted as a *house* that is modified by the word *brick*, which exhibits a PRODUCT-PRODUCER relationship between the two nouns in the compound. In contrast, the modifier and head in *house brick* exhibits a PART-WHOLE relationship, which is interpreted as a brick from a house, rather than the former interpretation of a house made of bricks. The set of SRs that we are concerned with in this paper is defined in Section 5.1.

Research on NCs can be categorised into four main groups: defining SRs, disambiguating the syntax of NCs, disambiguating the semantics of NCs, and interpreting NCs via SRs. Each task is detailed in Section 2.1. Interpreting NCs has received much attention of late, and the problem has been addressed in areas of machine translation (MT), information extraction (IE), and applications such as question-answering (QA). NCs pose a considerable challenge to computational linguistics due to the following issues (Lapata, 2002): (1) NCs are extremely productive; (2) the semantic relationship between the head noun and its modifier(s) is implicit; and (3) the interpretation of an NC can vary due to contextual and pragmatic factors. Due to these challenges, current NC interpretation methods are too error-prone to be employed directly in NLP applications without human intervention or preprocessing.

In this paper, we investigate the task of NC interpretation based on sense collocation. It has been shown that NCs with semantically similar components share the same SR (Kim and Baldwin, 2007); this is encapsulated by the phrase **sense collocation** in Moldovan et al. (2004). For example, *ap-*

ple pie has the same interpretation as *banana cake* of PRODUCT-PRODUCER. This can be predicted by the fact that the modifiers of both NCs (*apple* and *banana*, respectively) are semantically similar (they are both fruit), and the head nouns of both NCs (*pie* and *cake*, respectively) are a type of baked edible concoction. That the two NCs are based on the same combination of semantic classes is a strong predictor of the fact that they have the same SR.

One obvious problem when proceduralising semantic collocation in an interpretation method is data sparseness, i.e. we can't expect to find instances for all combinations of semantic classes, particularly if we have a rich inventory of semantic classes such as WORDNET. One approach to ameliorating the data spareness is bootstrapping, in the manner of Kim and Baldwin (2007), where new data is induced by substituting the components of the NCs with semantically similar terms. Our approach in this paper is to add related terms as features into a classifier. The related terms we add are the NC components' hypernyms, hyponyms, and sister words, based on the hypothesis that these related words can contribute to the disambiguation of SRs.

The remainder of this paper is structured as follows. In Section 2 we introduce previous research on NC interpretation and then talk specifically about the research directly relevant to this work. Section 3 and Section 4 describe our motivation and method, respectively. We describe the data used in our experiments in Section 5 and present the results of our experiments in Sections 6 and 7. Finally, we conclude our work in Section 8.

2 Related Work

This section presents a short description of the computational tasks relating to NCs, and then reviews research directly impinging on this research.

2.1 Background

The major tasks related to NCs involve syntactic and semantic disambiguation.

The first step in semantic disambiguation is the task of defining what relations exist in NCs. This has gained much attention in recent decades, as well as controversy, (Downing, 1977; Levi, 1979; Finin, 1980). In the study conducted by Levi (1979), it was claimed that there were 9 distinct SRs, which could be discretely defined and interpreted within NCs, while Finin (1980) claimed an unlimited number of SRs. The problems surrounding this task involve the issue of granularity versus coverage, which to date remains widely debated.

Syntactic disambiguation (called bracketing) is required when NCs are composed of more than 2 components, such as in the case of *computer science department*, introducing the need for phrasal disambiguation (Lauer, 1995; Nakov, 2005). Lauer (1995) proposed probabilistic models (based on dependency and adjacency analyses of the data). Later Nakov (2005) built upon this by adding linguistic features into these probabilistic models.

Methods employed in word sense disambiguation (WSD) have also been used to enhance NC interpretation; the noun components that comprise the NCs are disambiguated using these WSD techniques (Sparck Jones, 1983; Kim and Baldwin, 2007). Kim and Baldwin (2007) carried out experiments on automatically modeling WSD and attested the usefulness of conducting word sense analysis of an NC in determining its SR.

2.2 Previous Approaches to NC Interpretation

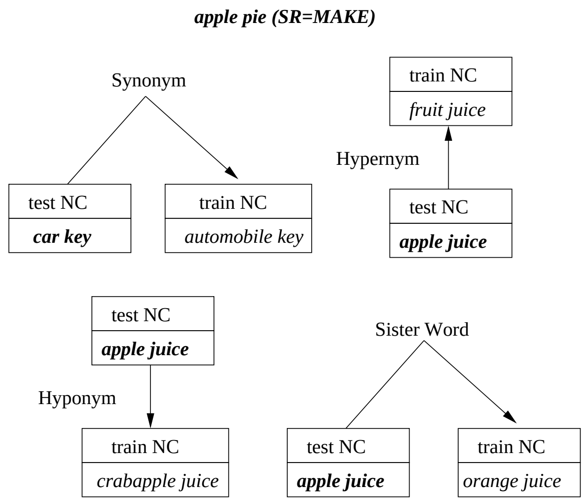
A majority of research undertaken in interpreting NCs has been based on two statistical methods: SEMANTIC SIMILARITY (Barker and Szpakowicz, 1998; Rosario, 2001; Moldovan et al., 2004; Kim and Baldwin, 2005; Nastase, 2006; Girju, 2007; Kim and Baldwin, 2007) and SEMANTIC INTERPRETABILITY (Vanderwende, 1994; Lapata, 2002; Kim and Baldwin, 2006; Nakov, 2006). Our work, based on an extension of the sense collocation approach, corresponds to the semantic similarity method.

A significant contribution to this area is by Moldovan et al. (2004), who used the sense collocation (i.e. pair of word senses) as their primary feature in disambiguating NCs. Many subsequent studies have been based on this sense collocation method, with the addition of other performance-improving features. For example, Girju (2007) added contextual information (e.g. the grammatical role and POS) and cross-lingual information from 5 European languages as features to her model. In contrast, Kim and Baldwin (2007) utilise sense collocations in a different way: instead of adding additional features in their model, they increase the size of their training data by substituting components of existing training instances to generate additional training instances (which is assumed to have the same SR as the original). For an SR to be preserved, the newly-generated NC must be semantically similar and hence maintain the same sense collocation as the original NC on which it was based. A number of researchers (Rosario (2001), Kim and Baldwin (2005), Nastase (2006), inter alia) have attempted to interpret NCs via implicit sense collocations. In particular, they have come up with various methods for avoiding direct WSD. Rosario (2001) retrieved the sense pairs in the context of a hierarchical class set for the biomedical domain. Kim and Baldwin (2005) used a word-level similarity measure to express the sense collocation of NCs. Nastase (2006) listed the hypernyms of components as sense features.

3 Motivation

As mentioned above, Moldovan et al. (2004) showed that the sense collocation of NCs is a key feature when interpreting NCs. Further research in this area has shown that not only synonymous NCs share the same SR, but NCs whose components are replaced with more loosely related words also commonly have the same SR as the original NCs (Kim and Baldwin, 2007). For example, *car factory*, *vehicle factory* and *truck factory*—corresponding to synonym, hypernym and sister word substitutions, respectively, over *automobile factory*—share the same SR of PRODUCT-PRODUCER as the source NC.

Figure 3 shows an example of semantic neighbours for the two NCs *car key* and *apple pie*. *Car key* can be interpreted as PRODUCT-PRODUCER by referring to the training NC *automobile key*, since they have the same sense collocation. With *apple juice*, the sense collocation method tries to locate matching sense collocations in the training data, and finds that *fruit juice* matches closely, with the modifier being a hypernym of *apple*. From this, we can hope to correctly interpret *apple juice* as having the SR PRODUCT-PRODUCER. In order to achieve this, we require some means of comparing nouns taxonomically, both vertically to capture hypernyms and hyponyms, and horizontally to capture sister words.



As intimated above, our motivation in conducting this research is to be able to include hypernym, hyponym and sister word information without using direct substitution over the training instances, but still preserving the essence of the sense collocation approach. The disadvantage of the method employed by Kim and Baldwin (2007) of recursively bootstrapping off a seed set of NCs via different lexical relations, is that noise will inevitably infect the training data, skewing the classifier performance. The original method described in Moldovan et al. (2004) only relies on observed sense collocations. The components of the NCs are represented as specific synsets in WORDNET, and the model does not capture related words. Hence, in this paper, we aim to develop a model that can take advantage of relatedness between WORDNET synsets via hypernyms, hyponyms and sister words, without the risk of losing semantic granularity or nintroducing noisy training data. Note that in Kim and Baldwin (2007), we used synonyms, hypernyms and sister words. As synonyms have an identical sense collocation within WORDNET (i.e. pairing of synsets) to the original NC, they are ignored in this research. Instead, we add hyponyms as a means of broadening the range of sense collocation.

4 Method

First, we describe the principal idea of the sense collocation approach to NC interpretation and the probability model proposed in Moldovan et al. (2004). Then we present our method using hypernyms, hyponyms and sister words in order to extend the sense collocation method.

4.1 Sense Collocation

The basic idea behind sense collocation in Moldovan et al. (2004) is based on the 'pair-of-word-senses' from the component nouns in NCs. They also introduced a probability model called **semantic scattering**, as detailed in Equations 1 and 2 below. In essence, the probability $P(r|f_i f_j)$ (simplified to $P(r|f_{ij})$) of a modifier and head noun with word sense f_i and f_j, respectively, occurring with SR r is calculated based on simple maximum likelihood estimation:

$$P(r|f_{ij}) = \frac{n(r, f_{ij})}{n(f_{ij})} \tag{1}$$

The preferred SR r^* for the given sense combination is that which maximizes the probability:

$$\begin{aligned} r^* &= \text{argmax}_{r \in R} P(r|f_{ij}) \\ &= \text{argmax}_{r \in R} P(f_{ij}|r)P(r) \end{aligned} \tag{2}$$

4.2 Adding Similar Words

We extend the approach of Moldovan et al. (2004) by adding similar words as features focusing on hypernyms, hyponyms and sister words of the modifier and head noun.

We accumulate the features for semantic relations based on different taxonomic relation types, from which we construct a feature vector to build a classifier over. The features of each taxonomic relation type are listed below. The first is features used in the original sense collocation method. The second, third and fourth are our experimental features, based on hypernyms, hyponyms and sister words respectively.

1. $< WS_{mod}, WS_{head} >$
2. $< WS_{mod}, H^i_{mod}, WS_{head}, H^i_{head} >$
3. $< WS_{mod}, O_{mod}, WS_{head}, O_{head} >$
4. $< WS_{mod}, S_{mod}, WS_{head}, S_{head} >$

where mod is the modifier, $head$ is the head noun, WS_{mod} is the WORDNET synset of the modifier, WS_{head} is the WORDNET synset of the head, H^i is an ith-degree ancestor (with direct hypernyms corresponding to H^1), O is a hyponym and S is a sister word. We include up to the 7th-degree ancestor (i.e. H^7), in line with the findings of Nastase (2006). Note that while a given synset has a unique hypernym in WORDNET (assuming no cycles, or the ability to remove cycles by precompiling a tree structure), it can have arbitrarily many hyponyms and sister words. Here, we take the cross product of the different hyponym and sister word candidates for a given synset.

We build our final classifier with TIMBL V6.0, a memory-based learner (Daelemans et al., 2004).

5 Data

Below, we outline the data used in our experiments.

5.1 Semantic Relation

The SR between a head and its modifier(s) in a NC tells us how to (default) interpret the NC. For example *door knob* corresponds to the PART-WHOLE relation, which means we can interpret *knob* as being part of a *door*. We sidestep the considerable challenge of developing an optimal set of semantic relation categories by using the set of SRs and data from the SEMEVAL-2007 nominal pair interpretation task (Girju et al., 2007). The SRs defined for the task are: CAUSE-EFFECT (CE), CONTENT-CONTAINER (CC), INSTRUMENT-AGENCY (IA), ORIGIN-ENTITY (OE), PART-WHOLE (PW), PRODUCT-PRODUCER (PP) and THEME-TOOL (TT). Table 1 provides a definition of each SR along with example NCs.

5.2 Data Collection

From the SEMEVAL-2007 annotated data (Girju et al., 2007), we collect two sets of data: a 2-class dataset and a 7-class dataset. The 2-class dataset is taken from the original SEMEVAL-2007 task, and comprises a set of positive and negative instances for each of the 7 SRs. The 7-class dataset is derived from this, by combining all positive NCs across the 7 SRs, in line with the methodology of Kim and Baldwin (to appear). The taxonomic relations are derived from WORDNET3.0. In each of the two sets, we use each of hypernyms, hyponyms and sister words. Table 2 shows the number of hyponyms and sister words in each dataset.

Semantic relation	Definition	Examples
Cause-Effect (**CE**)	N_1 is the cause of N_2	*virus flu, hormone growth, inhalation death*
Instrument-Agency (**IA**)	N_1 is the instrument of N_2, N_2 uses N_1	*laser printer, ax murderer, sump pump drainage*
Product-Producer (**PP**)	N_1 is a product of N_2, N_2 produces N_1	*honey bee, music clock, supercomputer business*
Origin-Entity (**OE**)	N_1 is the origin of N_2	*bacon grease, desert storm, peanut butter*
Theme-Tool (**TT**)	N_2 is intended for N_1	*reorganization process, copyright law, work force*
Part-Whole (**PW**)	N_1 is part of N_2	*table leg, daisy flower, tree forest*
Content-Container (**CC**)	N_1 is store or carried inside N_2	*apple basket, wine bottle, plane carge*

Table 1: The set of 7 semantic relations, where N_1 is the head noun and N_2 is a modifier

class	Hyponym		Sister word	
	mod	$head$	mod	$head$
7-classes	4866	4708	7167	7456
2-classes (CE)	1272	774	3220	2043
2-classes (IA)	955	1804	1726	3722
2-classes (PP)	1526	1688	3058	3009
2-classes (OE)	2394	1730	3861	2907
2-classes (TT)	1383	812	2767	1698
2-classes (PW)	1403	1770	2900	4117
2-classes (CC)	1598	820	2620	1909

Table 2: Total number of hyponym- and sister word-based NCs

6 7-way classification experiment

We ran our first experiment over the 7-class dataset. The baseline was computed using a *Zero-R* (i.e. majority class) classifier.[1] The performance of the original method proposed in Moldovan et al. (2004) is considered as a benchmark for our experiments. Table 3 shows the performance of the original sense collocation method and that of the extended sense collocation model proposed in this paper.

Table 3 shows that our method, combined with hypernyms, outperforms the original sense collocation method, with the highest accuracy of .588 achieved with 5th-degree ancestors of the head noun and modifier. This confirms that hypernyms are valuable in extending the range of sense collocation for NC interpretation.

In stark contrast to the results for hypernyms, the results for hyponyms and sister words significantly reduced the accuracy. The reason for this anomaly is that hypernyms are able to generalize the sense collocation without losing key discriminative features (i.e. the hypernyms always, by definition, subsume the original semantic information), while hyponyms and sister words add many sense collocations for which we have no direct evidence (i.e. we indiscriminately specialise the semantics without any motivation). Hence, hyponyms and sister words drastically blur the sense collocation.

The reason that the accuracy of the hypernym method drops in beyond a certain level is that the semantic collocations start to blend in together, and lose their power of discrimination.

7 2-way classification experiment

In our second experiment, we ran the systems over the original data from SEMEVAL-2007, in the form of a binary classifier for each of the 7 SRs. The performance of each of the 2-way classification tasks is shown in Table 4.

As we can see in Table 4, the basic pattern of the results is the same as for the 7-way classification task in Table 3. Adding hypernyms enhances performance, peaking for 4th-degree ancestors in this case at .679. As with the 7-way classification task, hyponyms and sister words degraded performance, for the same reasons as before.

Looking at the performance of each SR, we found that some SRs are easier to interpret than others. Notably, PRODUCT-PRODUCER and THEME-TOOL were high performers, while CAUSE-EFFECT was considerably harder to classify. These trends coincide with the system results for the SEMEVAL-2007 task. Girju et al. (2007) analyze this effect in terms of the intrinsic semantic complexity of the different SRs, and also the relative size of the training data for each SR. These effects are also observable in the breakdown of precision and recall of each SR in Figure 1.

As we used the data from the SEMEVAL-2007 task, we are able to directly compare the perfor-

[1] The majority class was PRODUCT-PRODUCER.

	B	M+	H^1	H^2	H^3	H^4	H^5	H^6	H^7	O	S
Accuracy	.217	.496	.544	.552	.573	.562	**.588**	.568	.557	.197	.142

Table 3: Results for the 7-way classification task: B = baseline, M+ = Moldovan et al. (2004) method, H^i = ith-order Hypernym, O = Hyponym and S = Sister word; the best performing system is indicated in **boldface**

	B	M+	H^1	H^2	H^3	H^4	H^5	H^6	H^7	O	S
CE	.547	.547	533	.573	.600	.606	.586	.607	**.630**	.467	.453
IA	.507	.581	.595	.608	.649	**.671**	.653	.629	.645	.500	.500
PP	.655	.667	.679	.691	.679	**.737**	.700	.690	.687	.655	.655
OE	.558	.636	.623	.610	.662	.645	.662	.625	**.712**	.558	.558
TT	.636	.697	.727	.712	.742	**.766**	.732	.717	.650	.515	.394
PW	.634	.620	.690	.690	.629	.657	.585	**.731**	.630	.633	.634
CC	.514	.676	**.703**	.689	.689	.676	.667	.647	.698	.446	.514
All	.579	.632	.649	.653	.662	**.679**	.654	.661	.667	.541	.534

Table 4: Results for each of the 2-way classification tasks: B = baseline, M+ = Moldovan et al. (2004) method, H^i = ith-order Hypernym, O = Hyponym and S = Sister word; the best performing system is indicated in **boldface**

mance of our method with the official results from the competing systems. Table 5 shows the three baselines provided by the SEMEVAL-2007 organisers (see Girju et al. (2007)). Here, *All True* is computed by guessing "true" for all relations, maximizing recall; *probability* is computed by randomly assigning "true" (or "false") with a probability matching the distribution of the labels in the training data for the given relation, and is intended to balance precision and recall; and *majority* is computed by assigning the majority class (either "true" or "false") from the training data for the given relation.

We also present the best-performing system and the average performance within group B from the SEMEVAL-2007 task (the grouping of systems which don't use gold-standard sense tags, and which also don't make use of the "query" used to source the examples).

As shown in Table 5, the performance of our method using hypernyms outperformed all three baselines. The performance using hyponyms and sister words only exceeded the All True and Probability baselines. The interesting point here is that although the method is meant for general-purpose interpretation not for the binary decision task, our proposed method with hypernyms achieves better results than the baselines and is competitive with the other systems in the original competition (average accuracy of group B = .656 vs. our best = .679). Therefore, we conclude that sense collocation integrated with hypernyms has the potential to extend the basic sense collocation method and improve performance for the NC interpretation task.

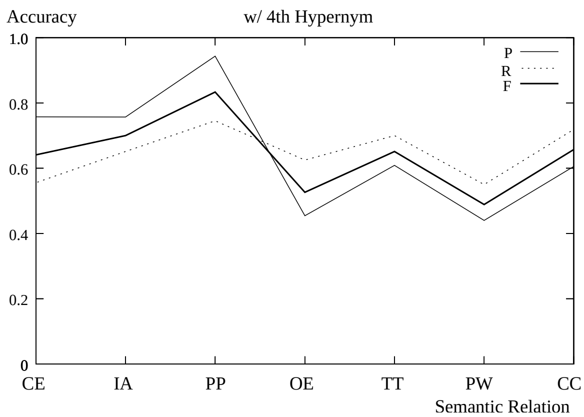


Figure 1: TPR for each of the binary tasks with 4th-degree hypernyms

8 Conclusion

In this paper, we have investigated the impact of using different taxonomic relations to expand a sense collocation method of NC interpretation. That is, we

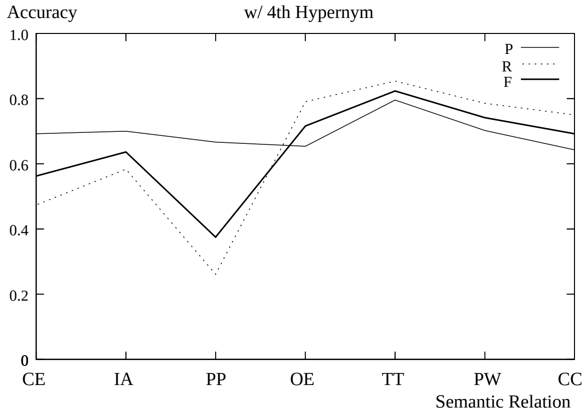


Figure 2: TNR for each of the binary tasks with 4th-degree hypernyms

Method	P	R	F	A
All True	.485	1.00	.648	.485
Probability	.485	.485	.485	.517
Majority	.813	.429	.308	.570
Best	.797	.698	.724	**.763**
Average	.650	.637	.631	.656

Table 5: Results of 2-way classification (P=precision, R=recall, F=F-score, A=accuracy)

experimented with the integration of similar terms into a sense collocation model. We added up to the 7th-degree hypernyms, direct hyponyms and direct sister words terms as features to the classifier. We ran experiments over 7-way and 2-way classification tasks using data from SEMEVAL-2007, and found that the inclusion of hypernym information significantly improved accuracy, while hyponyms and sister words degraded performance by arbitrarily over-specialising the sense information.

While intuitively all of hypernyms, hyponyms and sister words would appear to provide rich features for a sense collocation method, further research is needed to develop ways of successfully incorporating hyponyms and sister words into the NC interpretation task.

References

Ken Barker and Stan Szpakowicz. 1998. Semi-automatic recognition of noun modifier relationships. In *Proc. of the 17th International Conference on Computational Linguistics*, pp. 96–102, Montreal, Canada.

Walter Daelemans, Jakub Zavrel, Ko van der Sloot, and Antal van den Bosch. 2004. *TiMBL: Tilburg Memory Based Learner, version 5.1, Reference Guide*. ILK Technical Report 04-02.

Pamela Downing. 1977. On the Creation and Use of English Compound Nouns. *Language*, 53(4):810–842.

Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, USA.

Timothy W. Finin. 1980. *The Semantic Interpretation of Compound Nominals*. Ph.D. thesis, University of Illinois.

Roxana Girju. 2007. Improving the interpretation of noun phrases with cross-linguistic information. In *Proc. of the 45th Annual Meeting of the Association of Computational Linguistics*, pp. 568–575, Prague, Czech Republic.

Roxana Girju, Preslav Nakov, Vivi Nastase, Stan Szpakowicz, Peter Turneyand Deniz Yuret. 2007. SemEval-2007 task 04: classification of semantic relations between nominals. In *Proc. of Semantic Evaluation Workshop*, Prague, Czech Republic.

Su Nam Kim and Timothy Baldwin. 2005. Automatic interpretation of noun compounds using WordNet similarity. In *Proc. of the 2nd International Joint Conference On Natural Language Processing*, pp. 945–956, JeJu, Korea.

Su Nam Kim and Timothy Baldwin. 2006. Interpreting semantic relations in noun compounds via verb semantics. In *Proc. of the 44th Annual Meeting of the Association for Computational Linguistics and 21st International Conference on Computational Linguistics (COLING/ACL-2006)*. pp. 491–498, Sydney, Australia.

Su Nam Kim and Timothy Baldwin. 2007. Disambiguating noun compounds. In *Proc. of the 22nd AAAI conference on Artificial Intelligence (AAAI-07)*, 2007, pp.901-906, British Columbia, Canada.

Su Nam Kim and Timothy Baldwin. 2007. Interpreting noun compounds using bootstrapping and sense collocation. In *Proc. of the Pacific Association for Computational Linguistics (PACLING)*, Melbourne, Australia.

Su Nam Kim and Timothy Baldwin. to appear. Benchmarking noun compound interpretation. In *Proc. of the 3rd International Joint Conference on Natural Language Processing (IJCNLP-08)*, Hyderabad, India.

Maria Lapata. 2002. The disambiguation of nominalizations. *Computational Linguistics*, 28(3):357–388.

Mark Lauer. Designing Statistical Language Learners: Experiments on Noun Compounds, Ph.D thesis. 1995, Macquarie University.

Judith Levi. 1979. The syntax and semantics of complex nominals. In *The Syntax and Semantics of Complex Nominals*. New York:Academic Press.

Dan Moldovan, Adriana Badulescu, Marta Tatu, Daniel Antohe, and Roxana Girju. 2004. Models for the semantic classification of noun phrases. In *Proc. of the HLT-NAACL 2004 Workshop on Computational Lexical Semantics*, pp. 60–67, Boston, USA.

Preslav Nakov and Marti Hearst. 2005. Search engine statistics beyond the n-gram: Application to noun compound bracketing. In *Proc. of the 9th Conference on Computational Natural Language Learning (CoNLL-2005)*, pp. 17-24, Ann Arbor, Michigan, USA.

Preslav Nakov and Marti Hearst. 2006. Using verbs to characterize noun-noun relations. In *Proc. of the 12th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA)*, Bularia.

Vivi Nastase and Jelber Sayyad-Shiarabad and Marina Sokolova and Stan Szpakowicz. 2006. Learning noun-modifier semantic relations with corpus-based and WordNet-based features. In *Proc. of the 21st Annual Conference on Artificial Intelligence (AAAI-06)*, pp. 781-787, Boston, MA, USA.

Barbara Rosario and Hearst Marti. 2001. Classifying the semantic relations in noun compounds via a domain-specific lexical hierarchy. In *In Proc. of EMNLP*, 82–90

Karen Sparck Jones. 1983. Compound noun interpretation problems. *Computer Speech Processing*, Prentice-Hall, Englewood Cliffes, NJ, USA.

Diarmuid Seaghdha and Ann Copestake. 2007. Co-occurrence contexts for noun compound interpretation. In *Proc. of the ACL-2007 Workshop on A Broader Perspective on Multiword Expressions*, Prague, Czech Republic.

Lucy Vanderwende. 1994. Algorithm for automatic interpretation of noun sequences. In *Proc. of the 15th Conference on Computational linguistics*, pp. 782–788.