

# Mixing WordNet, VerbNet and PropBank for studying verb relations

Maria Teresa Pazienza\*, Marco Pennacchiotti\*, Fabio Massimo Zanzotto†

\*University of Roma Tor Vergata  
Via del Politecnico 1,  
Roma, Italy  
{pazienza, pennacchiotti}@info.uniroma2.it  
† DISCo, University of Milano Bicocca  
Via B. Arcimboldi 8  
Milano, Italy  
zanzotto@disco.unimib.it

## Abstract

In this paper we present a novel resource for studying the semantics of verb relations. The resource is created by mixing sense relational knowledge enclosed in WordNet, frame knowledge enclosed in VerbNet and corpus knowledge enclosed in PropBank. As a result, a set of about 1000 *frame pairs* is made available. A *frame pair* represents a pair of verbs in a peculiar semantic relation accompanied with specific information, such as: the syntactic-semantic frames of the two verbs, the mapping among their thematic roles and a set of textual examples extracted from the PennTreeBank. We specifically focus on four relation: Troponymy, Causation, Entailment and Antonymy. The different steps required for the mapping are described in detail and statistics on resource mutual coverage are reported. We also propose a practical use of the resource for the task of Textual Entailment acquisition and for Question Answering. A first attempt for automate the mapping among verb arguments is also presented: early experiments shows that simple techniques can achieve good results, up to 85% F-Measure.

## 1. Introduction

The study of verb syntax and semantics is the main focus of many NLP researches, as it is of great help in supporting a large area of natural language applications (Question Answering, Information Extraction, etc.).

Indeed, verbs play a central role in semantics, as they convey the core meaning of sentences: the situation described in a sentence is in fact expressed through its main verb. All other components of the sentence usually depend and turn around the main verb. Nouns and noun phrases express the participants to the situation, while adjectives and other grammatical elements are used to better specify and describe the situation and its participants.

To better understand the importance of verbs syntactic and semantic behaviour for NLP applications, consider a classical QA system which has to answer the following user question:

*"What country does Israel fear for its nuclear ability?"*

A possible answer could be derived through the use of linguistic knowledge and inference reasoning, from a textual fragment like:

*"Iran scares Israel with its nuclear ability."*

In order to carry out such a complex matching, the system should firstly understand the question. It thus must have some knowledge about the syntactic-semantic behaviour of the verb *fear*, that is, the relation between the syntax and the semantics of the verb arguments. It should then identify the subject of the question as the semantic *Experiencer* of the situation, the object as the *Theme*, and the argument introduced by the preposition *in* as the semantic *Predicate*. On the other hand, the system should also recognize the syntactic construction of the retrieved snippet governed by

*scare*. Here, the subject of the sentence is the *Cause*, while the object plays the role of the *Experiencer* and the argument introduced by *with* is an *Oblique* semantic role.

Furthermore, to match question and answer, the system should have some knowledge about verb relations (see Fig. 1): specifically, it should know that the verb *scare* entails the verb *fear* (one is scared by something only if he fears it). Finally, the system must know how to match verb arguments. In particular, considering that the *Experiencer* of the two verbs are the same ("*Israel*"), and that the *Predicate* of the question matches the *Cause* of the answer, it should be able to infer that the *Theme* searched in the question is actually the *Cause* in the snippet ("*Iran*"). The answer is thus straightforward:

*"Israel fears Iran for its nuclear ability."*

Such kind of linguistic inferences are of great use not only in QA, but also in many other NLP tasks (Paraphrasing, Machine Translation, Textual Entailment, etc.) but are still far from being fully exploited. In fact, even if dedicated syntactic-semantic resources (as VerbNet and PropBank for verb structures, WordNet for word senses) and corpora (e.g., PennTreeBank) are already available, what still lack is a consistent and beneficial integration among them. Specifically, a tight integration between semantic resources and corpora would allow the use of statistical and Machine Learning techniques for supporting tasks on verb semantics.

In this paper we propose a simple but effective way to integrate these resources, by creating a mapping among WordNet verb relations, VerbNet verb frames and the PennTreeBank corpus, using the PropBank database as a bridge between the linguistic resources and the corpus. We thus aim to exploit the principle of linking theory, as implemented in

the existing resources, to connect relational verb semantics to syntax and finally to surface realizations.

The goal of our proposed mapping is twofold. From the one hand, the link among verb sense relations to their frames and eventually to their instances in a corpus can be exploited to carry out complex inference reasoning as described in the previous example. From the other hand we want to verify the mutual coverage of the different resources, as it is a fair indicator for both estimating the usefulness of such kind of mappings and for verifying how language usage (PropBank) adhere to a specific theory (VerbNet).

The paper is organized as follows. In Sec.2. we briefly describe the different resources involved in the process. In Sec.3. we outline related works on resource mapping. In Sec.4. we then describe our mapping process, together with some statistics on resources coverage. Finally, in Sec.5. we propose an example of effective usage of the mapping for Textual Entailment patterns discovery.

## 2. WordNet, VerbNet, PropBank and PennTreeBank

### 2.1. WordNet

WordNet (Miller, 1995) is a lexical database aiming at modelling the human linguistic knowledge in a coherent repository inspired by psycholinguistic theories. WordNet covers approximately 150.000 words, among nouns, verbs, adjectives and adverbs. Words are organized in *synsets*, each representing a lexicalized concept that is linguistically represented by a set of synonymy words and a gloss describing the synset itself. WordNet also represents semantic links among synsets (such as troponymy, entailment, antonymy, etc.), thus structuring the lexicon in a network of synset relations.

Verbs are an important part of the network. Roughly 11.000 verbs are present, divided in 24.632 senses: the polisemy of verbs is thus quite high, as each verbs has in average 2,3 senses. Verbs can be linked through specific relations, such as troponymy, antonymy, entailment and causation. These relations express (directely or indirectly) a lexical entailment relation between two verbs (see Sec.5.).

Unluckily, as it is mainly devoted to model semantics, WordNet lacks information about verb syntax, which is of primary importance in creating a complete resource for working on verb behaviours.

### 2.2. VerbNet

VerbNet (Kipper et al., 2000) is a hierarchical verb lexicon based on Levin's verb classes representing verbs syntactic and semantic information. The main idea is that the syntactic frames of a verb reflect their semantics: verbs sharing similar syntactic behaviour can be then clustered in semantically coherent classes. Each verb class is constituted by the a set of verbs, their shared *syntactic frames*, *thematic roles* and *selectional restrictions*. Moreover, semantic predicates are added to each class to better describe its semantic behaviours.

VerbNet 2.0 contains 237 hierarchically organized classes (mainly inspired by Levin's hierarchy) and 5000 verbs. Each verb in a class is semantically unambiguous and is

explicitly linked to the WordNet synset(s) representing its sense. As WordNet verb synset lexicon is more fine-grained than VerbNet classes, more than one synsets can be linked to a verb in a class. In average each verb in VerbNet is mapped to 3,2 WordNet senses. Moreover, a verb appearing in two different classes usually refers to a different sense (and thus links to a different synset).

VerbNet provides an effective and useful link between the syntax and the semantics of a verb. Yet, as the representation of complex verb relations and the connection to a reference corpus are not the focus of VerbNet, a mapping to other resources (e.g. WordNet, PropBank) can be usefully exploited.

### 2.3. PropBank and PennTreeBank

PropBank (Babko-Malaya et al., 2004) is formed by a verb lexicon and a semantically annotated corpus. The lexicon contains about 3600 verbs. Each verb is represented by a *frame*. Each frame is composed by one or more *framesets*, that refer to specific verb senses. In all, PropBank contains 5050 framesets. Each frameset is accompanied with its set of semantic roles (*roleset*), identified by generic argument labels (Arg0, Arg1, ..., ArgM). The mapping between roles and labels is roleset specific, that is, a label is usually assigned to different roles across the lexicon.

Semantic roles in PropBank are more specific than thematic roles in VerbNet. Indeed, while in VerbNet roles are general and valid across different classes, in PropBank they are strictly tied to a specific roleset. As a consequence, VerbNet has only 20 thematic roles, while PropBank has more than 1400 roles. Each PropBank roleset is mapped, when possible, to a VerbNet class, as it will be described in Sec.4..

As in VerbNet, verb senses in PropBank are fairly coarse-grained with respect to WordNet, as they are derived by studying the frequency of the verb frame structures in the corpus. Frameset are in fact created grouping those verb syntactic frames that share the same semantic roles.

PropBank frames are used to semantically annotate the Penn Wall Street Journal Treebank II with predicate-argument structures: each occurrence of a verb is properly annotated together with its argument labels. In all, more than 110.000 PropBank instances are annotated. The final corpus and the lexicon can be then used as a resource for learning useful linguistic phenomena related to verb semantics.

Yet, what is missed in PropBank is a link to WordNet: it is thus not possible to directly use PropBank knowledge in conjunction with WordNet information on verb relations.

## 3. Related Work

The mapping and the integration of different syntactic-semantic resources is today a main concern. Many studies have been thus devoted to this problem. In (Kipper et al., 2002), (Kipper et al., 2004) a mapping between VerbNet and PropBank is proposed. Finding links between the two resources is not an easy task as it seems. Indeed, VerbNet classes and PropBank framesets have been created using different methodologies: VerbNet is havily based on the Levin's classes theory, while PropBank was primarily built looking at verb usage in the PennTree corpus. The mutual

coverage of the two resources is thus a useful means to understand to what extent language theory can fit language usage in a specific framework. In (Kipper et al., 2004) each PropBank frameset has been manually linked, when possible, to the VerbNet class expressing its syntactic-semantic behaviour. Each PropBank role is then mapped to the corresponding VerbNet thematic roles.

Not all framesets in PropBank have a link to VerbNet, as a verb sense can be present in PropBank but not in VerbNet. Moreover, not all roles have a mapping to a thematic role. A PropBank frameset can be mapped to more than one VerbNet class, as PropBank senses are often more coarse-grained.

VerbNet coverage seems reasonable: 78.62% of PropBank sentences in the corpus have an *exact matching* (all roles are mapped) to a VerbNet class. Notwithstanding, some VerbNet frames have no corpus instantiations. No mention is made in the study about mutual coverage of the two lexicon (i.e., the set of frames and classes).

Among other studies, (Shi and Mihalcea, 2005) propose a mapping between VerbNet classes and FrameNet frames and between VerbNet roles selectional restrictions and specific WordNet semantic classes. The study aims at building a unified resource to support semantic parsing.

In (Giuglea and Moschitti, 2004) the link between VerbNet and PropBank proposed in (Kipper et al., 2002) is used together with a semi-automatic mapping from VerbNet to FrameNet to improve performance of a role labelling system.

Yet, to our knowledge, no specific research has been devoted so far to the mapping (and the exploitation) of WordNet relations, VerbNet classes and PropBank lexicon and corpus.

#### 4. Mapping resources

Our goal is to link WordNet 2.0 (verb sense relational knowledge) to VerbNet 2.0 (verb sense frame knowledge) and finally to the PropBank corpus (verb sense frame repository). The final objective is thus to have a large set of linguistic examples of verb pairs that have some semantic relation and specific predicate-argument structures. Once such integrated resource is available, it will be possible to study and automatically learn how the predicate-argument structures of two verbs are related using the set of corpus examples.

For example in WordNet the verbs *scare* and *fear* are in entailment relation. Specifically, the first sense of *scare* entails the second sense of *fear* ( $scare_1 \Rightarrow fear_2$ ). The VerbNet classes corresponding to the two verbs are respectively *amuse-31.1* and *admire-31.2*. It is then possible to extract the syntax frames for the two verbs, as reported in Fig.1. Moreover, the two verb classes are mapped in PropBank to the frameset *scare.01* and *fear.01*. Thanks to this mapping it is thus possible to extract from the PropBank corpus all the sentences related to the frameset. At the end of the mapping process it is thus available a large corpus of sentences corresponding to the predicate structure allowed for the verb sense relation  $scare_1 \Rightarrow fear_2$ . Such a resource can be then refined to find interesting interaction between the two verb senses structure. For instance, it

can be learned from the example sentences (see Sec.5.) the mapping among the verbs thematic roles (e.g., the *Cause* of *scare* is mapped to the *Theme* of *fear*).

The mapping among WordNet, VerbNet and PropBank is not straightforward as it seems, for two main reasons:

- the three resources were built independently and with different goals and methodologies. WordNet stems from a psycholinguistic theory and its aim is to model a mental lexicon. VerbNet, as a reflection of Levin's classes, is an instantiation of a linguistic theory: the goal is to model a coherent syntactic-semantic interface. PropBank frameset system is a lexicon basically built from a corpus: it thus stems from language usage and not from language theory. As a gap generally exists between language theory and usage, and between a psycholinguistic and a purely linguistic representation of lexicon, it is obvious that main discrepancies can occur in the mapping process.
- The grain of the resources is different. WordNet is fine-grained with respect to verbs and verb senses, while VerbNet and PropBank are both coarse-grained but not aligned one with the other. Thus, it is not always straightforward to obtain a coherent mapping between verb senses. Moreover, the mismatch between the grain of VerbNet thematic roles and PropBank rolesets is another major challenge (as described in (Kipper et al., 2002)).

All these issues must be taken into account when mapping the different resources, as described hereafter.

##### 4.1. WordNet to VerbNet

The mapping between WordNet and VerbNet verb senses is straightforward, as WordNet senses are explicitly represented in VerbNet classes. Over the 24.632 verb senses in WordNet only 4.712 have a correspondence in VerbNet (19.2%). The low semantic coverage of VerbNet is due to the scope of the two resources: WordNet aim at covering the whole lexicon, while VerbNet looks only at verbs with specific behaviours. On the other hand, all VerbNet senses are present in WordNet. As VerbNet senses are coarse-grained, they can be mapped to more than one synset.

This verbe sense mapping is used to find pairs of VerbNet verb predicate structures corresponding to WordNet verb sense relations. Four type of relations (Miller, 1995) have been studied : troponymy, entailment, causation (generically called *lexical entailment* relations) ant antonymy. WordNet accounts for 71.364 verb pairs in these relations (e.g., *win*  $\rightarrow$  *compete*). Due to the low VerbNet coverage only 11.015 relations (15.4%) can be mapped in VerbNet (i.e., both verbs are present in VerbNet); the number is further reduced to 8.964 when pairs with verbs in the same VerbNet class are discarded. Obviously this is a strong limitation for the final resource, as many relational knowledge in WordNet is not used. Notwithstanding it is a good starting point for implementing automatic methodologies for mapping.

**Frame Pair:** *scare - fear 01*

Relation : causation

WordNet Senses : scare1 - fear2

VerbNet Classes : amuse-31.1 - admire 31.2

VerbNet Frame mapping :

1f. Cause scare Experiencer

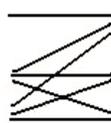
2f. Cause scare Experiencer Prep(with) Oblique

3f. Cause [+genitive] Oblique scare Experiencer

1h. Experiencer fear Theme

2h. Experiencer fear Theme Prep(for) Predicate

3h. Experiencer Predicate Prep(in) Theme



Argument Mapping:

Experiencer - Experiencer

Cause - Theme

Oblique - Predicate

Figure 1: The *frame pair* for verbs *scare-fear*.

## 4.2. VerbNet to PropBank

The mapping between VerbNet and PropBank is obtained looking at the explicit link in PropBank framesets to VerbNet classes. Over the 4.498 PropBank framesets (verb senses), 1.969 have a *direct* mapping to VerbNet (43,8%), i.e. they refer to a class in which the verb explicitly appears. On the other hand, 2.363 over 5000 VerbNet verb senses are mapped to PropBank (47.3%). Roughly, PropBank seems thus to have a better coverage than VerbNet. As a final step thematic roles and PropBank roles are also manually mapped.

## 4.3. Frame pair resource

Our final resource mapping WordNet, VerbNet and PropBank thus consists of a set of *frame pairs*. A *frame pair* is composed by:

- A pair of verb senses in a specific semantic relation;
- The syntactic-semantic behaviours of the two verbs, as expressed by VerbNet framesets.
- A set of mappings between the frames of the two verbs for which the semantic relation holds (*VerbNet Frame mapping*);
- A mapping between the thematic roles for the two verbs (*Argument mapping*);
- A set of textual examples for each pair, derived from the PennTreeBank. The actual sentences extracted from the PennTree are assigned to a verb in VerbNet recomposing the sentence using the argument structure of the syntactic frame in the VerbNet class, as done in (Kipper et al., 2004).

In all, our resource contains 989 *frame pairs* (see Table 1). In average, each *frame pair* has 50 PennTreeBank example sentences. An example of *frame pair* is shown in Fig.1

*VerbNet Frame mapping* and *Argument mapping* are currently on work and done by hand. In Sec.5. possible techniques for automate these processes are presented.

Many strategies could be applied in the different phases to improve the coverage of the final resource. For example VerbNet verb coverage with respect to WordNet could be improved augmenting classes with new verbs imported from PropBank through an *indirect mapping*. Many PropBank verbs are in fact linked to VerbNet classes where they are not already present. In such cases, during the building process we discarded the link, as there is not explicit connection between the new verbs and their WordNet synset. However we are working at a possible disambiguation process assigning WordNet senses to PropBank verbs, without using VerbNet as a disambiguation resource.

## 5. Textual Entailment Patterns Discovery

A first application of our resource is the discovery of textual entailment patterns. A Textual Entailment patterns, as defined in (Dagan and Glickman, 2004), is formed by a *text* template  $T$  and an *hypothesis* template  $H$ .  $T$  and  $H$  are two language expressions such that  $T$  entails  $H$ , accompanied with syntactic properties and possible free slots. For example, the pattern:

$$X_{subj} scare Y_{obj} Z_{with} \rightarrow Y_{subj} fear X_{obj} Z_{for}$$

states that for any syntactically coherent instantiation of  $X$ ,  $Y$  and  $Z$ , entailment holds.

*Entailment pattern acquisition* is the task of collecting these generalized forms, using different techniques ranging from statistical counts to linguistic analysis. Once a

Relation	WordNet	WordNet-VerbNet	WordNet-VerbNet-PropBank
Troponymy	66.756	8.124	888
Entailment	2.251	468	49
Causation	1.278	150	26
Antinomy	1.079	222	26
All	71.364	8.964	989

Table 1: Number of verb relations present in WordNet that can be mapped to VerbNet and successively to PropBank.

large collection of entailment patterns has been acquired, it can be straightforwardly used to retrieve entailment relations in new texts, as needed for specific applications. In such a way, a QA system could map a question like “*What country does Israel fear for its nuclear ability?*” into the  $H$  template  $Israel_{subj} fear Y_{obj} nuclear\_ability_{for}$ . Then, all the  $T$  templates related to  $H$  could be retrieved from the collection. Finally, all the answers could be extracted from a corpus, as surface forms matching the  $T$  template (e.g.  $Iran_{subj} scare Israel_{obj} nuclear\_ability_{with}$ ).

Our aim is to use the frame pairs obtained by the mapping described in Sec.4. to infer such entailment patterns. Indeed, given two related verb senses, it is possible to retrieve through VerbNet all their syntactic frames. Once these syntactic frames are available, the problem is then to find which frames of the two verbs are in entailment relation, and how the arguments must be matched (*VerbNet Frame mapping and Argument mapping*).

Argument mapping is in theory a difficult task, as in VerbNet thematic roles are not meant to be consistent with the syntactic inversion that can take place between two related verbs. In particular this assumption is valid for *causation* verbs. For these class of verbs the subject of the entailed verb  $H$  is usually the object of the entailing verb (e.g.  $X_{subj} raise Y_{obj} \rightarrow Y_{subj} rise$ ). In other terms, in most cases the subject of  $T$  carries out an action that changes the state of the object, that is then described by the verb  $H$ . In such cases, VerbNet roles are not always consistent. For example, in  $X_{Agent} raise Y_{Patient} \rightarrow Y_{Agent} rise$ , the *Patient* of *raise* becomes the *Agent* of *rise*.

Yet, we experimented the consistency of VerbNet thematic roles, in order to verify if the argument matching task could be carried out by simply keeping the roles across the relation (e.g.  $Agent_T \rightarrow Agent_H$ ) (*simple mapping*). Surprisingly, we found that this technique shows good results. We randomly selected 15 verb relations among the lexical entailments of our resource and we manually build a gold standard, mapping the arguments of the  $T$  verbs to the arguments of the  $H$  verbs. We then estimated Precision and Recall of the *simple mapping* technique, by counting how many roles of the  $T$  verb were correctly mapped to the roles of the  $H$  verb. Simple mapping shows a Precision of 96,4% and a Recall of 66,7% (*F-measure* is 78,7%). Low Recall indicates that in many cases an argument of  $T$  is not present in  $H$ . While results must be taken as only a first exploratory study, they however seem to outline a certain degree of role consistency among VerbNet classes, even for many causation verbs. For example in :

$$X_{Experiencer} hurt Y_{Patient} \rightarrow Y_{Patient} ache$$

the *Patient* role is consistent across the two verbs.

Yet, more complex methods can be used for supporting the simple mapping technique to improve recall, that is particularly low for causation verbs (only 60%). For example the mapping can be boosted by looking at the context of the arguments: all the examples of each frame are retrieved from the PennTree corpus, using the mapping between VerbNet and PropBank. Once examples for all arguments of the two templates are available methods based on Distributional Hypothesis ((Harris, 1985)) can be applied. As the Distributional Hypothesis states that words with similar meaning tend to appear in similar context, we can extend the idea to verb arguments. We can state that two matching thematic roles are filled with similar terms. It is then possible to use *co-occurrence vectors* (Pantel, 2005) to estimate similarity between two argument contexts. Firstly, co-occurrence vector for each argument  $a$  of the templates classes are built as a list of fillers  $f$  accompanied with their frequency or their pointwise mutual information ( $mi_{af}$ ).  $mi_{af}$  is defined as follows:

$$mi_{af} = \log \frac{P(a,f)}{P(f) \times P(a)}$$

where  $P(f)$  is estimated as the frequency of the filler in the class,  $P(a)$  the frequency of the classes argument and  $P(a, f)$  the frequency of the filler in that class argument. The similarity between each argument  $a_{Ti}$  in  $T$  and each argument  $a_{Hj}$  in  $H$  is then computed as the distance between the corresponding co-occurrence vectors. The distance  $D(a_{Ti}, a_{Hj})$  is estimated using cosine similarity (Salton and McGill, 1983), as:

$$D(a_{Ti}, a_{Hj}) = \frac{\sum_f a_{Ti} \times a_{Hj}}{\sqrt{\sum_f a_{Ti}^2 \times \sum_f a_{Hj}^2}}$$

We applied this technique by extracting all nouns present in the fillers and using them to build the co-occurrence vectors. Early results show a slightly lower precision/recall than the *simple mapping* technique.

Notwithstanding, the two methods can be used together to improve recall. We then applied the co-occurrence techniques to map those arguments that are not matched by *simple mapping*. On the set of 15 relations, this mixed technique shows an improvement in F-measure from 78,7% to 85,5% (Recall improves from 66,7% to 88%, at cost of a lower Precision, 84%).

## 6. Conclusions

In this paper we presented a mapping among WordNet, VerbNet and PropBank for studying semantic verb relations. Our main goal was to find the verb frames and textual

examples of verbs in a specific semantic relation stated in WordNet. While the mutual coverage of the resources is not very high, we were able to infer almost 1000 *frame pairs*. We believe that this repository could be successfully used to support NLP applications, where verb semantics is a concern, such as in QA, as it has been showed throughout the paper. Yet, in order to exploit this knowledge, automatic or semi-automatic techniques are needed to map arguments across the frames. We showed that such mapping is feasible and that simple techniques boosted with corpus-based methods should guarantee good performance. However, an extensive experiment is still required together with the application of more refined methods based on the Distributional Hypothesis.

As a second goal, we verified that the mutual coverage between VerbNet and PropBank is fairly good. That seems to indicate that Levin's verb theory instantiated in VerbNet has an effective correspondence in language usage, as expressed in the PennTreeBank corpus.

As a future work we are interested in extending the experiments for argument mapping and in using different techniques for this task, such as those based on the selectional restrictions, as the one proposed in (Shi and Mihalcea, 2005).

We also plan to use the discovered *frame pairs* to infer textual entailment patterns. Once available, these patterns will be used both to give a deeper inside on the textual entailment phenomenon and to support QA applications. Moreover, the discovered *frame pairs* represent a valuable resource that could be used to build Machine Learning models for studying verb relation semantics.

## 7. References

- O. Babko-Malaya, M. Palmer, N. Xue, A. Joshi, and S. Kulick. 2004. Proposition bank ii: Delving deeper. In A. Meyers, editor, *HLT-NAACL 2004 Workshop: Frontiers in Corpus Annotation*, pages 17–23, Boston, Massachusetts, USA, May 2 - May 7. Association for Computational Linguistics.
- Ido Dagan and Oren Glickman. 2004. Probabilistic textual entailment: Generic applied modeling of language variability. In *Learning Methods for Text Understanding and Mining*, Grenoble, France.
- Ana-Maria Giuglea and Alessandro Moschitti. 2004. Knowledge discovery using framenet, verbnet and propbank. In A. Meyers, editor, *Workshop on Ontology and Knowledge Discovering at ECML 2004*, Pisa, Italy.
- Zellig Harris. 1985. Distributional structure. In Katz J.J., editor, *The Philosophy of Linguistics*, pages 26–47. Oxford University Press, New York.
- Karin Kipper, Hoa Trang Dang, and Martha Palmer. 2000. Class-based construction of a verb lexicon. pages 691–696.
- Karin Kipper, Martha Palmer, and Owen Rambow. 2002. Extending propbank with verbnet semantic predicates. In *AMTA 2002 Workshop: Workshop on Applied Interlinguas*, Tiburon, CA, USA.
- Karin Kipper, Benjamin Snyder, and Martha Palmer. 2004. Extending a verb-lexicon using a semantically annotated corpus. In *urth International Conference on Language Resources and Evaluation (LREC 2004)*, Lisbon, Portugal.
- George A. Miller. 1995. WordNet: A lexical database for English. *Communications of the ACM*, 38(11):39–41, November.
- Patrick Pantel. 2005. Inducing ontological co-occurrence vectors. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*, pages 125–132, Ann Arbor, Michigan, June. Association for Computational Linguistics.
- G Salton and M J McGill. 1983. *Introduction to Modern Information Retrieval*. McGraw-Hill.
- Lei Shi and Rada Mihalcea. 2005. Putting pieces together: Combining framenet, verbnet and wordnet for robust semantic parsing. In A. Meyers, editor, *6th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing-05)*, Mexico City, Mexico, May 2 - May 7.