A Large-Scale Extension of VerbNet with Novel Verb Classes

Karin Kipper
Computer and Information Science Department
University of Pennsylvania
kipper@linc.cis.upenn.edu, nryant@unagi.upenn.edu

Anna Korhonen
Computer Laboratory
University of Cambridge, UK
alk23@cl.cam.ac.uk

Neville Ryant and Martha Palmer
Department of Linguistics
University of Colorado at Boulder
martha.palmer@colorado.edu

Abstract
Lexical classifications have proved useful in supporting various linguistic and natural language processing (NLP) tasks. The largest verb classification in English is Levin's (1993) work. VerbNet (Kipper-Schuler 2006) – the largest computational verb lexicon currently available for English – provides detailed syntactic-semantic descriptions of Levin classes. While the classes included are extensive enough for some NLP use, they are not comprehensive. Korhonen and Briscoe (2004) have proposed a significant extension of Levin's classification which incorporates 57 novel classes for verbs not covered (comprehensively) by Levin. Korhonen and Ryant (2005) have recently supplemented this with another extension including 53 additional classes. This paper describes the integration of these two extensions into VerbNet. The result is an extensive Levin style classification for English verbs which now provides over 90% token coverage of the PropBank data, making possible the use of supervised machine learning.

1 Introduction

Lexical classes, defined in terms of shared meaning components and similar (morpho-) syntactic behavior of words (Jackendoff, 1990; Levin, 1993), have attracted considerable interest in both linguistics and natural language processing (NLP). These classes are useful for their ability to capture generalizations about a range of (cross-)linguistic properties.

For example, verbs which share the meaning component of 'manner of motion' (such as *travel, run, walk*), behave similarly also in terms of subcategorization (*I traveled/ran/walked, I traveled/ran/walked to London, I traveled/ran/walked five miles*) and usually have zero-related nominals (*a run, a walk*). Although the correspondence between the syntax and semantics of words is not perfect and these classes do not provide means for full semantic inferencing, their predictive power is nevertheless considerable.

NLP systems can benefit from lexical classes in a number of ways. Such classes define the mapping from the surface realization of arguments to predicate-argument structure, and are therefore an important component of any system which needs the latter. As the classes approximate higher level abstractions (e.g., syntactic or semantic features) they can be used as a principled means for abstracting away from individual words when required. Their predictive power can help compensate for the lack of sufficient data that fully exemplifies the behavior of relevant words. Lexical classes have proved helpful in supporting a number of tasks, such as computational lexicography (Kipper et al., 2000), machine translation (Dorr, 1997), word sense disambiguation (Prescher et al., 2000) and subcategorization acquisition (Korhonen, 2002). While this work has met with success, it has so far been small in scale. Large-scale exploitation of the classes in real-world tasks or in highly domain-sensitive tasks (e.g. information extraction) has not been possible because no comprehensive lexical classification is available.

VerbNet (VN) (Kipper-Schuler 2006)[1] is the largest on-line verb lexicon currently available for English. It provides detailed syntactic-semantic descriptions of Levin classes organized into a refined taxonomy. While the original version of VN has proved useful for a variety of natural language tasks (Swift, 2005; Hensman and Dunnion 2004; Crouch and Holloway, 2005; Swier and Stevenson, 2004), it mainly dealt with verbs taking noun (NP) and prepositional phrase (PP) complements and thus suffered from limited coverage. Some experiments have been reported which indicate that it should be possible, in the future, to automatically supplement VN with new classes and member verbs from corpus data (Brew and Schulte im Walde, 2002; Korhonen et al., 2003; Kingsbury 2004). Such automatic approaches can greatly reduce manual classification and enable quick tuning of the resource to specific applications. However, the very development of these approaches requires access to a target classification more extensive than that available currently.

VerbNet has recently been integrated with 57 new classes from Korhonen and Briscoe's (2004) (K&B) proposed extension to Levin's original classification (Kipper et al., 2006). This work has involved associating detailed syntactic-semantic descriptions to the K&B classes, as well as organizing them appropriately into the existing VN taxonomy. In the present paper we first provide a brief summary of this recent extension and then introduce a novel extension: the incorporation of an additional set of 53 new classes from Korhonen and Ryant (2005) (K&R) into VN. The outcome is a freely available resource which constitutes the most comprehensive and versatile Levin-style verb classification for English. After the two extensions VN has now also increased our coverage of PropBank tokens (Palmer et. al., 2005) from 78.45% to 90.86%, making feasible the creation of a substantial training corpus annotated with VN thematic role labels and class membership assignments, to be released in 2007. This will finally enable large-scale experimentation on the utility of syntax-based

[1] http://verbs.colorado.edu/~kipper/verbnet.html

classes for improving the performance of syntactic parsers and semantic role labelers on new domains.

This paper is organized as follows. VerbNet is briefly described in Section 2. The classes of K&B and K&R are presented in Section 3. Section 4 describes the integration of these classes into VN. Finally, Section 5 describes how this integration affected VerbNet and its coverage, and discusses on-going and future work.

2 VerbNet

VerbNet is a hierarchical domain-independent, broad-coverage verb lexicon with mappings to other lexical resources such as WordNet (Miller, 1990; Fellbaum, 1998), Xtag (XTAG Research Group, 2001), and FrameNet (Baker et al., 1998). VerbNet is organized into verb classes extending Levin's original classification through refinement and addition of subclasses to achieve syntactic and semantic coherence among members of a class. Each verb class in VN is completely described by thematic roles, selectional restrictions on the arguments, and frames consisting of a syntactic description and semantic predicates with a temporal function, in a manner similar to the event decomposition of Moens and Steedman (1988).

Each VN class contains a set of syntactic descriptions, or syntactic frames, depicting the possible surface realizations of the argument structure for constructions such as transitive, intransitive, prepositional phrases, resultatives, and a large set of diathesis alternations. Semantic restrictions (such as animate, human, organization) are used to constrain the types of thematic roles allowed by the arguments, and further restrictions may be imposed to indicate the syntactic nature of the constituent likely to be associated with the thematic role. Syntactic frames may also be constrained in terms of which prepositions are allowed.

Each frame is associated with explicit semantic information, expressed as a conjunction of boolean semantic predicates such as 'motion,' 'contact,' or 'cause.' Each semantic predicate is associated with an event variable *E* that allows predicates to specify when in the event the predicate is true (*start*(*E*) for preparatory stage, *during*(*E*) for the culmination stage, and *end*(*E*) for the consequent stage). Figure 1. shows a complete entry for a frame in VerbNet class *Hit-18.1*.

Class ***Hit-18.1***			
Roles and Restrictions: Agent[+int_control] Patient[+concrete] Instrument[+concrete]			
Members: bang, bash, hit, kick, ...			
Frames:			
Name	Example	Syntax	Semantics
Basic Transitive	Paula hit the ball	Agent V Patient	cause(Agent, E) manner(during(E), directedmotion, Agent) !contact(during(E), Agent, Patient) manner(end(E),forceful, Agent) contact(end(E), Agent, Patient)

Figure 1. Simplified VerbNet entry for Hit-18.1 class

2.3 Status of VN

Before integrating the proposed classes, VN had descriptions for over 4,600 verb senses (over 3,400 lemmas) distributed in 191 first-level classes, and 74 new subclasses. The information in the lexicon has proved useful for various NLP tasks such as word sense disambiguation and semantic role labeling. In the original VN Levin's taxonomy has gained considerably in depth, but not in breadth. Verbs taking adjectival (ADJP), adverbial (ADVP), particle, predicative, control and sentential complements were still largely excluded, except where they showed interesting behavior with respect to NP and PP complementation. In Section 5, we show how this integration greatly increased our coverage of corpus data.

3 Description of the new classes

3.1 K&B Classes

The resource of Korhonen and Briscoe (2004) includes a substantial extension to Levin's classification with 57 novel classes for verbs as well as 106 new diathesis alternations. The classes were created using the following semi-automatic approach:[2]

Step 1: A set of diathesis alternations were constructed for verbs not covered extensively by Levin. This was done by considering possible alternations between pairs of subcategorization frames (SCFs) in the comprehensive classification of Briscoe (2000) which incorporates 163 SCFs (a superset of those listed in the ANLT (Boguraev et al., 1987) and COMLEX Syntax dictionaries (Grishman et al., 1994)), focusing in particular on those SCFs not covered by Levin. The SCFs define mappings from surface arguments to predicate-argument structure for bounded dependency constructions, but abstract over specific particles and prepositions. 106 new alternations were identified manually, using criteria similar to Levin's.

Step 2: 102 candidate lexical-semantic classes were selected for the verbs from linguistic resources of a suitable style and granularity: (Rudanko, 1996; Rudanko, 2000), (Sager, 1981), (Levin, 1993) and the LCS database (Dorr, 2001).

Step 3: Each candidate class was evaluated by examining sets of SCFs taken by its member verbs in syntax dictionaries (e.g., COMLEX) and whether these SCFs could be related in terms of diathesis alternations (106 novel ones or Levin's original ones). Where one or several alternations were found which captured the sense in question, a new verb class was created.

Steps 1-2 were done automatically and step 3 manually. Identifying relevant alternations helped to identify additional SCFs, which often led to the discovery of additional alternations. For those candidate classes which had an insufficient number of member verbs, new

[2] See Korhonen and Briscoe (2004) for the details of this approach and http://www.cl.cam.ac.uk/users/alk23/classes/ for the latest version of the classification.

members were searched for in WordNet. These were frequently found among the synonyms, troponyms, hypernyms, coordinate terms and/or antonyms of the extant member verbs. The SCFs and alternations discovered during the identification process were used to create the syntactic-semantic description of each novel class.

For example, a new class was created for verbs such as *order* and *require*, which share the approximate meaning of "direct somebody to do something". This class was assigned the following description (where the SCFs are indicated by number codes from Briscoe's (2000) classification):

Order Verbs
SCF 57: *John ordered him to be nice*
SCF 104: *John ordered that he should be nice*
SCF 106: *John ordered that he be nice*
Alternating SCFs: 57↔104, 104↔106

The work resulted in accepting, rejecting, combining and refining the 102 candidate classes and – as a by-product – identifying 5 new classes not included in any of the resources used. In the end, 57 new verb classes were formed, each associated with 2-45 member verbs. Table 1 shows a small sample of these classes along with example verbs. The evaluation of the novel classes showed that they can be used to support an NLP task and that the extended classification has a good coverage of the English verb lexicon.

Class	Example Verbs
URGE	*ask, persuade*
FORCE	*manipulate, pressure*
WISH	*hope, expect*
ALLOW	*allow, permit*
FORBID	*prohibit, ban*
HELP	*aid, assist*

Table 1. *Examples of K&B's classes*

3.2 K&R Classes

While working on VerbNet and the K&B classes, Korhonen and Ryant (2005) discovered 53 additional verb classes which cover a wide range of different complements (including, but not limited to, sentential complements). The classes were identified using the same methodology as in 3.1 (Step 3), associated with 2-37 member verbs and assigned similar syntactic descriptions as K&B classes. Table 2. presents a small sample of these classes along with member verbs.

Class	Example Verbs
INTERROGATE	*interrogate, question*
ADJUST	*adjust, adapt*
SUBJUGATE	*suppress, subdue*
BEG	*request, supplicate*
COMPREHEND	*grasp, comprehend*

Table 2. Examples of K&R's classes

4 Integrating the resources

Although the new classes of K&B and K&R are similar in style to the Levin classes included in VN, their integration to VN proved a major task. The first step was to assign the classes VN-style detailed syntactic-semantic descriptions. This was not straightforward because the new classes lacked explicit semantic descriptions and had syntactic descriptions not directly compatible with VN's descriptions.

Also some of the descriptions available in VN had to be enriched for the new classes. The second step was to incorporate the classes into VN. This was complicated by the fact that K&B and K&R are inconsistent in terms of granularity: some classes are broad while others are fine-grained. Also the comparison of the new classes to Levin's original classes had to be done on a class-by-class basis: some classes are entirely new, some are subclasses of existing classes, while others require reorganization of original Levin classes. These steps, which were conducted largely manually in order to obtain a reliable result, are described in sections 4.1 and 4.2, respectively.

4.1 Syntactic-Semantic Descriptions of Classes

Assigning syntactic-semantic descriptions to the new classes involved work on both VN and the new classifications. The different sets of SCFs used by the resources required creating new roles, syntactic descriptions and restrictions on VN. The set of SCFs in K&B and K&R is broad in coverage and relies, in many cases, on finer-grained treatment of sentential complementation than present in VN. Therefore, VN's syntactic descriptions had to be enriched with a more detailed treatment of sentential complementation. On the other hand, prepositional SCFs in K&B and K&R do not provide VN with explicit lists of allowed prepositions as required, so these had to be created and added to the classes. Also, no syntactic description of the surface realization of the frames was included in K&B and K&R and had to be created. Finally, information about semantic (thematic) roles and restrictions on the arguments was created from scratch and added to K&B and K&R.

4.1.1 Changes

In integrating the new classes, it was found that none of the 21 VN thematic roles seemed to appropriately convey the semantics of the arguments for some classes. Two new thematic roles were added to VN, *Content* and *Proposition* in order to make this integration.

Not all VN's syntactic frames had a counterpart in Briscoe's classification. This discrepancy is the by-product of differences in the design of the two resources. Briscoe abstracts over prepositions and particles whereas VN differentiates between otherwise identical frames based on the precise types of prepositions that a given class of verbs subcategorizes for. Additionally, VN may distinguish two syntactic frames depending on thematic roles. But regarding sentential complements the opposite occurs, with VN conflating SCFs that Briscoe's classification considers distinct. In integrating the proposed classes into VN it was necessary to greatly enrich the set of possible syntactic restrictions VN allows on clauses. The new set of possible syntactic restrictions consists of 55 features accounting for object control, subject control, and different types of complementation.

Integrating the new classes also required enriching VN's set of semantic predicates. Whenever possible, existing VN predicates were reused. However, as many of the incoming classes represent concepts entirely novel to VN, it was necessary to introduce new predicates to adequately provide descriptions of the semantics of incoming classes.

4.2 Integrating the K&B Classes into VerbNet

The initial set of classes proposed by K&B featured 57 classes. Of these, two were rejected as being either insufficiently semantically homogeneous or too small to be added to the lexicon, with the remaining 55 selected for incorporation. The classes fell into three different categories regarding Levin's classification: 1) classes that could be subclasses of existing Levin classes; 2) classes that require a reorganization of Levin classes; 3) entirely new classes.[3]

A total of 42 classes fell into the first category. Although some of these overlapped to an extent with existing VN classes semantically the syntactic behavior of their members was distinctive enough to allow them to be added as new classes (35 novel classes were actually added as new classes while 7 others were added as new subclasses).

A total of 13 classes fell into the second category. These classes overlapped significantly in some way with existing VN classes (either too close semantically or syntactically) and required restructuring of VN. This integration was done in one of two ways: either by merging the proposed classes with the related VN class (e.g., WANT and Want-32.1, PAY and Give-13.1); or by adding the proposed class as a novel class but making modifications to existing VN classes (this was the case for some classes of *Verbs with Predicative Complements*, which classify more naturally in terms of sentential rather than NP or PP complements)...

For further details of the integration of the K&B classes into VN see Kipper et al. (2006).

4.3 Integrating the K&R classes into VerbNet

Integrating the second set of candidate classes proceeded much as detailed in 4.1 and 4.2. Each of the 53 candidate classes was assigned a VN-style class description before being evaluated for inclusion. Of the 53 suggested classes, 7 *were* omitted as they did not fully meet the requirements of Levin style syntactic-semantic classes, 11 were decided to overlap to a reasonable extent with a pre-existing class, and 36 were added as new classes (1 candidate class was divided into 2 new classes).

4.3.1 Novel Classes and Subclasses

In total, 35 classes from K&R were regarded as sufficiently novel for addition to VN without restructuring of an existing VN class. In addition, one class was divided into 2 new classes, PROMISE and ENSURE. As with K&B, 10 classes overlapped semantically, but not syntactically with existing VN classes, and hence were added as new subclasses. Examples

---

[3] Levin focused mainly on NP and PP complements, but many verbs classify more naturally in terms of sentential complementation.

of such classes include the proposed classes INTERROGATE and BEG, which were added as subclasses of the classes concerning *Communication*. The remaining 26 candidate classes were added as new classes. Examples include the classes REQUIRE, DOMINATE, SUBJUGATE, and HIRE, all of which express novel concepts.

4.3.2 Additions to existing classes

11 of the candidate classes overlapped significantly both syntactically and semantically with an existing class. Examples include CLARIFY (overlaps the EXPLAIN class of the first candidate set), DELEGATING_POWER (overlaps ALLOW of first candidate set), BEING_IN_CHARGE_OF (overlaps second candidate set DOMINATE). Unlike with K&B classes, very little restructuring was needed for these cases. In each of the 11 cases, the proposed class contained a subset of the SCFs in the class it overlapped with or contained one or two additional SCFs which were compatible with the pre-existing class.

4.4 Comparison of K&B and K&R

4.4.1 New subcategorization frames

Three K&R classes required the use of new SCFs not appearing in either VN or any of the classes of the first candidate set to give their full syntactic description: USE, BASE, and SEEM. A total of 4 new SCFs were used, distributed among these classes, listed below in Table 3.

Class	*SCF*	*Example*
USE	NP-P-POSSING	I used the money for Mary's leaving the country.
BASE	NP-P-WH-S	They based their claim on whether he happened to mention the danger.
BASE	NP-P-NP-ING	They based their objections on him failing to mention the dangers.
SEEM	ADJ-PRED-RS	He appeared crazy.
SEEM	ADJ-PRED-RS	He appeared well.

Table 3. *Examples of subcategorization frames for K&R classes*

4.4.2 *Differences in content between K&B and K&R*

The most salient difference among the two candidate sets is in the categories of activities they include. Many of the 42 classes of the K&B set tended to cluster among 3 broad categories:

i) classes describing the interaction of two animate entities:

There are 14 classes which describe interactions or relationships among entities in some social context (see Examples a), b), and c)). The interaction can be either cooperative or non-cooperative and the two entities may or may not be thought to exist in some power relationship.

(a) FORCE
"John forced Bill to go home."

(b) CONSPIRE
"John conspired with Bill to overthrow the government."
(c) BATTLE
"John battled with Bill over the insult."

ii) classes describing the degree of engagement of an entity with an activity:
There are 11 classes that involve an agent and an activity in which the agent is involved, but differ in how the Agent approaches the activity (see Examples a), b), and c)).
(a) TRY
"John tries to keep the house clean."
(b) NEGLECT
"John neglected to wash the car."
(c) FOCUS
"John focused on getting the car clean."

iii) classes describing the relation of an entity and some abstract idea:
There are 6 classes that describe relations between and abstract entities, such as whether the idea is a novel contribution of the entity or the entity's attitude toward the idea.
(a) DISCOVER
"John discovered that he can hold his breath for two minutes."
(b) WISH
"John wishes to go home."

In contrast, the classes of the K&R set seem to address a much broader range of concepts (note that they also cover a wider range of complementation pattern types than the K&B classes). There, is, again, a group of 10 classes that could broadly be related as describing social interactions among animate entities (i.e., DOMINATE, SUBJUGATE, HIRE). The remaining of the classes tend to form small clusters of 2-4 classes, or are among the 10 completely idiosyncratic classes.

i) Small clusters:
For example, both ESTABLISH and PATENT classes describe activities of bringing into existence, but, unlike the existing *Create-26.4* verbs, these new classes relate to the creation of abstractions such as organizations or ideas.
(a) ESTABLISH
"John tries to keep the house clean."
(b) PATENT
"I patented my discovery with a gleeful smile."

ii) Idiosyncratic classes:
Examples of these include classes such as USE, SEEM, and MULTIPLY.
(a) USE
"I utilized the new methodology in my research."

(b) SEEM
"John seems a fool."

(c) MULTIPLY
"The children divided each sum by the number of items in a simple exercise of statistics."

5 Conclusion and Future Work

Integrating the two recent extensions to Levin classes into VerbNet was an important step in order to address a major limitation of Levin's verb classification, namely the fact that verbs taking ADJP, ADVP, predicative, control and sentential complements were not included or addressed in depth in that work. This limitation excludes many verbs that are highly frequent in language. A summary of how this integration affected VN and the result of the extended VN is shown in Table 4. The figures show that our work enriched and expanded VN considerably. The number of first-level classes grew significantly (from 191 to 274), there was also a significant increase in the number of verb senses and lemmas, along with the set of semantic predicates and the syntactic restrictions on sentential complements.

An obvious question from the NLP point of view is the practical usefulness of the extended VN. Korhonen and Briscoe (2004) showed that the K&B classes now incorporated in VN can be used to significantly aid an NLP task (subcategorization acquisition) and that VN extended with the K&B classes has a good coverage over the English verb lexicon as evaluated against WordNet. When evaluating the usefulness of the current VN (extended with both K&B and K&R), the key issue is coverage, given the insufficient coverage has been the main limitation of the use of verb classes in practical NLP so far. In order to address this question, we investigated the coverage of the current VN over PropBank (Palmer et al., 2005) – the annotation of the Penn Treebank II with dependency structures. The list of verbs in VN before the class extensions included 3,445 lemmas (unique verb senses) which matched 78.45% of the verb tokens in the annotated PropBank data (88,584 occurrences). The new version of VN extended with verbs and classes included in K&B and K&R contains 3,769 lemmas. This greatly increased the coverage of VN to now match 90.86% of the PropBank verb occurrences (102,600 occurrences).

Such an extensive resource makes it feasible to create substantial training corpus annotated with VN thematic role labels and class membership assignments. This resource, which we plan release in 2007, will finally enable large-scale experimentation on the utility of syntax-based classes for improving the performance of syntactic parsers and semantic role labelers on new domains.

	VN without K&B classes	Extended VN
First-level classes	191	274
Thematic roles	21	23
Semantic predicates	64	94
Syntactic restrictions (on sentential compl)	3	55
Number of verb senses	4656	5257
Number of lemmas	3445	3769

Table 4. Summary of the Lexicon's Extension

Currently we are investigating ways to further extend VN's coverage with new classes and particularly with new member verbs. We already started searching for additional members using automatic methods, e.g. clustering (Kingsbury and Kipper, 2003; Kingsbury 2004, Korhonen et al., 2003). We are also considering to include, in the future, in VN statistical information concerning the relative likelihood of different classes, SCFs and alternations for verbs in corpus data, using, for example the automatic methods proposed by McCarthy (2001) and Korhonen (2002). Such information can be highly useful for statistical NLP systems utilizing lexical classes.

Acknowledgements

This work was supported by National Science Foundation Grants NSF-9800658, VerbNet, NSF9910603, ISLE, NSF-0415923, WSD, the DTO-AQUAINT NBCHC040036 grant under the Univ. of Illinois subcontract to Univ. of Pennsylvania 2003-07911-01 and DARPA grant N66001-00-1-8915 at the Univ. of Pennsylvania. Any opinions, findings, and conclusions or recommendations expressed here are those of the authors and do not necessarily reflect the views of the National Science Foundation, DARPA, or the DTO.

References

Baker, C. F., Fillmore, C. J., Lowe, J. B. (1998), 'The Berkeley FrameNet project', in *Proceedings of the 17th International Conference on Computational Linguistics (COLING/ACL-98)*, Montreal, pp. 86-90.

Boguraev, B., Briscole, T., Carroll, J., Carter, D., Grover, C. (1987), 'The derivation of a grammatically-indexed lexicon from the Longman Dictionary of Contemporary English', in *Proceedings of the* 25th *annual meeting of ACL*, pp. 193-200, Stanford, CA.

Brew, C., Schulte im Walde, S (2002), 'Spectral clustering for German verbs', in *Conference on Empirical Methods in Natural Language Processing*, Philadelphia, USA.

Briscoe, T. (2000), *Dictionary and System Subcategorisation Code Mappings*. Unpublished manuscript, http://www.cl.cam.ac.uk/users/alk23/subcat/subcat.html, University of Cambridge Computer Laboratory.

Crouch, D., Holloway King, T. (2005), 'Unifying lexical resources', in *Proceedings of Interdisciplinary Workshop on the Identification and Representation of Verb Features and Verb Classes*, Saarbruecken.

Trang Dang, H. (2004), *Investigations into the Role of Lexical Semantics in Word Sense Disambiguation*. Ph.D. thesis, CIS, University of Pennsylvania. Dorr. B. J. (1997), Large-scale dictionary construction for foreign language tutoring and interlingual machine translation. *Machine Translation*, 12(4), pp. 271-325.

Dorr, B. J. (1997), Large-scale dictionary construction for foreign language tutoring and interlingual machine translation. *Machine Translation*, 12(4), pp. 271-325.

Dorr, B. J. (2001), 'LCS Verb Database', in *Online Software Database of Lexical Conceptual Structures and Documentation*, University of Maryland.

Fellbaum, C. (ed.) (1998), *WordNet: An Eletronic Lexical Database*. Language, Speech and Communications, Cambridge, Massachusetts, MIT Press.

Grishman, R., Macleod, C., Meyers, A. (1994), 'COMLEX syntax: building a computational lexicon', in *Proceedings of the International Conference on Computational Linguistics*, Kyoto.

Hensman, S., Dunnion, J. (2004), 'Automatically building conceptual graphs using VerbNet and WordNet', in *Proceedings of the 3rd International Symposium on Information and Communication Technologies (ISICT)*, pp. 115-120, Las Vegas.

Jackendoff, R. (1990), *Semantic Structures*, Cambridge, Massachusetts, MIT Press.

Kingsbury, P., Kipper, K. (2003), 'Deriving Verb-Meaning Clusters from Syntactic Structure', in *Workshop on Text Meaning*, Edmonton.

Kingsbury, P. (2004), 'Verb clusters from PropBank annotation', in *Technical Report*, University of Pennsylvania, Philadelphia.

Kipper K., Trang Dang, H., Palmer, M. (2000), 'Class-based construction of a verb lexicon', in *AAAI/IAAI*, pp. 691-696.

Kipper-Schuler, K. (2005), *VerbNet: A broad-coverage, comprehensive verb lexicon*. PhD. Thesis. Computer and Information Science Dept., University of Pennsylvania. Philadelphia.

Kipper, K., Korhonen, A., Ryant, N., Palmer, M. (2006), 'Extending VerbNet with Novel Classes'. To appear in *Proceedings of the 5th International Conference on Language Resources and Evaluation*, Genoa.

Korhonen A., Briscoe, T. (2004), 'Extended Lexical-Semantic Classification of English Verbs', in *Proceedings of the HLT/NAACL Workshop on Computational Lexical Semantics*, Boston.

Korhonen, A., Krymolowski, Y., Marx, Z. (2003), 'Clustering polysemic subcategorization frame distributions semantically', in *Proceedings of the 41st Annual Meeting of ACL*, Sapporo, pp. 64-71.

Korhonen, A. (2002) 'Semantically motivated subcategorization acquisition', in *ACL Workshop on Unsupervised Lexical Acquisition*, Philadelphia.

Korhonen, A., Ryant, N. (2005), *53 Novel Lexical-Semantic Verb Classes*. Unpublished manuscript http://www.cl.cam.ac.uk/users/alk23/classes2/.

Levin, B. (1993), *English Verb Classes and Alternation, A Preliminary Investigation*. The University of Chicago Press.

McCarthy, D. (2001), *Lexical Acquisition at the Syntax-Semantics Interface: Diathesis Alternations, Subcategorization Frames and Selectional Preferences*. Ph.D. thesis, University of Sussex.

Miller, G. A. (1990) 'WordNet: An on-line lexical database', *International Journal of Lexicography*, 3(4), pp. 235-312.

Moens, M., Steedman, M. (1988), 'Temporal Ontology and Temporal Reference', *Computational Linguistics*, 14, pp. 15-38.

Palmer, M., Gildea, D., Kingsbury, P. (2005), 'The Proposition Bank: A Corpus Annotated with Semantic Roles', Computational Linguistics Journal, 31(1).

Prescher, D., Riezler, S., Rooth, M. (2000), 'Using a probabilistic class-based lexicon for lexical ambiguity resolution', in *18th International Conference on Computational Linguistics*, Saarbrucken, pp. 649-655.

Rudanko, J. (1996), *Prepositions and Complement Clauses*, Albany, State University of New York Press.

Rudanko, J. (2000), *Corpora and Complementation*, University Press of America.

Sager, N. (1981), *Natural Language Information Processing: A Computer Grammar of English and Its Applications*,Addison-Wesley Publising Company, MA.

Shi, L., Mihalcea, R. (2005), 'Putting pieces together: Combining FrameNet, VerbNet and WordNet for robust semantic parsing', in *Proceedings of the Sixth International Conference on Intelligent Text Processing and Computational Linguistics*, Mexico City.

Swier, R., Stevenson, S. (2004), 'Unsupervised semantic role labelling', in *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona, pp. 95-102.

Swift, M. (2005), 'Towards automatic verb acquisition from VerbNet for spoken dialog processing', in *Proceedings of Interdisciplinary Workshop on the Identification and Representation of Verb Features and Verb Classes*, Saarbruecken.

XTAG Research Group (2001), 'A lexicalized tree adjoining grammar for English', Technical Report IRCS-01-03, IRCS, University of Pennsylvania.