PropBank: Semantics of New Predicate Types

Claire Bonial[1,2], Julia Bonn[2], Kathryn Conger[2], Jena D. Hwang [1,2], Martha Palmer[1,2]

[1]Department of Linguistics, University of Colorado at Boulder
[2]Institute of Cognitive Science, University of Colorado at Boulder
Bonial, Julia.Bonn, Kathryn.Conger, Hwangd, Martha.Palmer @colorado.edu

Abstract

This research focuses on expanding PropBank, a corpus annotated with predicate argument structures, with new predicate types; namely, noun, adjective and complex predicates, such as Light Verb Constructions. This effort is in part inspired by a sister project to PropBank, the Abstract Meaning Representation project, which also attempts to capture "who is doing what to whom" in a sentence, but does so in a way that abstracts away from syntactic structures. For example, alternate realizations of a *destroying* event in the form of either the verb *destroy* or the noun *destruction* would receive the same Abstract Meaning Representation. In order for PropBank to reach the same level of coverage and continue to serve as the bedrock for Abstract Meaning Representation, predicate types other than verbs, which have previously gone without annotation, must be annotated. This research describes the challenges therein, including the development of new annotation practices that walk the line between abstracting away from language-particular syntactic facts to explore deeper semantics, and maintaining the connection between semantics and syntactic structures that has proven to be very valuable for PropBank as a corpus of training data for Natural Language Processing applications.

Keywords: Predicate semantics, Semantic role labelling, Syntax, Natural Language Processing

1. Introduction

The annotated corpus PropBank (Palmer et al., 2005) represents an ongoing effort to provide the information necessary to map between the syntactic analysis of a sentence and the conceptual structure of an event relation. Previously, the annotation effort has focused on event relations expressed solely by verbs. (A separate but related effort, NomBank, focused on nouns (Meyers et al., 2004).) However, a complete representation of event relations within and across sentences requires expanding that focus to additional syntactic realizations of the same eventuality, including expressions in the form of nouns, adjectives and multi-word expressions. Capturing the semantics of these additional predicates has presented challenges unique to each predicate type, as an attempt is made to assign semantic roles to all arguments and adjuncts of a predicate, but the syntactic environment in which these arguments and adjuncts are realized can be very different. This research discusses how these challenges were addressed to successfully expand the PropBank corpus: first by developing guidelines specific to the annotation of each predicate type, but also by developing practices that will eventually allow annotations to focus more on semantics alone (concepts and relations), while moving beyond language-particular syntactic facts. This new direction is in part inspired by a desire for greater interoperability with the Abstract Meaning Representation (AMR) project (Banarescu et al., 2013). A primary goal of AMR is to provide training data for meaning-based machine translation; therefore, a deliberate effort is made to focus on representing semantics in a language-independent fashion.

2. PropBank and Other Lexical Resources

There are currently five English lexical resources that provide explicit semantic role labels for use in data annotation: FrameNet, VerbNet, LIRICS, EngVallex and Prop-Bank. These resources have been created independently and with differing goals, but they are surprisingly compatible. They differ primarily in the granularity of the semantic role labels. PropBank uses very generic labels such as Arg0 and Arg1,[1] as in:

1. President Bush has approved duty-free treatment for imports of certain types of watches.

   Relation (REL): approved
   Arg0: President Bush
   Arg1: duty-free treatment for imports of certain types of watches.

EngVallex uses non-numbered labels (e.g. ACT (Actor), PAT (Patient), ADDR (Addressee), ORIG (Origin) and EFF (Effect)), which, with the exception of the first two, make them more descriptive, irrespective of the verb of which they are the argument. In addition to providing several alternative syntactic frames and a set of semantic predicates, VerbNet marks the PropBank Arg0 as an Agent, and the Arg1 as a Theme. FrameNet labels them Grantor and Action respectively, and puts them in the Grant Permission frame. The additional semantic richness provided by Verb-Net and FrameNet does not contradict PropBank, and can be seen as complementary. The LIRICS project, Linguistic InfRastructure for Interoperable ResourCes and Systems, has made a serious study of these different frameworks and of the theoretical linguistics background, resulting in a detailed set of Semantic Role definitions.[2] Within the LIRICS framework, Arg0 and Arg1 would be labeled Agent and Theme, respectively, like VerbNet.

---

[1]The other numbered arguments in PropBank, Arg2-5, are quite verb-specific.

[2]http://let.uvt.nl/general/people/bunt/docs/
LIRICS_semrole.htm

Again, note that a clear difference in these resources is how fine-grained the role labels are. PropBank has what can be considered very coarse-grained numerical labels, but these correspond to very fine-grained verb-specific labels found in the PropBank lexicon. FrameNet uses very fine-grained labels, specific to the semantics of a particular domain. VerbNet and LIRICS represent an intermediate level of granularity. Research in automatic semantic role labeling has demonstrated the importance of the level of granularity of semantic roles: Yi et al. (2007) and Loper et al. (2007) both demonstrate that because VN labels are more generalizable across verbs than PropBank labels, they are easier for semantic role labeling systems to learn; however, Merlo and van der Plas (2009) found that the differing levels of granularity of PropBank and VN were both useful, and therefore suggest complementary use of both resources.

Using these resources together can also overcome coverage limitations of any single corpus. SemLink (Palmer, 2009), an ongoing effort to map PropBank, VerbNet, FrameNet, and the OntoNotes sense groupings (which, in turn, map to WordNet senses (Fellbaum et al., 1998)), facilitates use of these resources together. For example, in their attempts to use FrameNet as automatic semantic role labeling training data, Giuglea and Moschitti (2006) found that they obtain better results by interconnecting FrameNet to VerbNet and PropBank, thereby overcoming gaps in FrameNet's coverage. Similarly, Shi and Mihalcea (2005) found that they could build an improved resource for semantic parsing by linking FrameNet, VerbNet, and Wordnet.

3. PropBank Background

PropBank is somewhat distinct from these resources in that PropBank annotation was developed specifically to provide training data for supervised machine learning classifiers. It provides semantic information, including the basic "who is doing what to whom" much like the other resources, but the definitions are in the form of predicate-by-predicate semantic role assignments. PropBank annotation firstly consists of the selection of a 'roleset,'[3] or a coarse-grained sense of the predicate, which includes a listing of the roles, expressed as generic argument numbers, associated with that sense. Here, for example, is the roleset for the verb *fear*:

Arg0: entity afraid
Arg1: afraid of what?

These argument numbers, along with a variety of modifier tags, such as temporal and locative, are assigned to natural language sentences drawn from a variety of corpora. The goal is to assign these simple, general-purpose labels consistently across the many possible syntactic realizations of the same event participant or semantic role.

As training data, PropBank provides valuable information that allows for mapping between syntactic structures and semantic roles; a process that can be more complex than

it initially seems. For example, in the following two sentences,

2. The flame melted the wax.

3. The wax melted.

a standard syntactic parser represents *the wax* as the verb's direct object in the first sentence and its subject in the second. There is nothing overt to indicate that it has the same conceptual relation in both cases despite the fact that it is expressed syntactically in a different way. We can capture this by annotating *the wax* as having the same semantic role (or conceptual relation) in both sentences. It would typically be labelled the Patient, or Arg1 in the case of PropBank, the participant undergoing a change of state. Note that both sentences are in the active voice, and not the passive voice. In *The wax was melted by the flame*, the passive provides syntactic evidence that *the wax* is playing the same role (Patient) that it plays in example 2. Since the particular pair of syntactic variations illustrated by *melt* does not occur with every transitive verb, it is not easily predictable. By providing manually annotated information on the ways in which a semantic role can be associated with different syntactic realizations of the same verb, PropBank comprises training data allowing for automatic systems to map back and forth between syntax and semantics.

What sets PropBank apart from FrameNet and VerbNet is in that its Frame Files and their rolesets are specifically tailored to the usage of the predicate in naturally occurring corpus data. FrameNet provides semantically rich conceptual frames, defining all potential frame elements that can be instantiated in a specified event, and VerbNet generalizes verb semantic behavior based on syntactic alternations. PropBank contributes usage-based formulations of verb-argument structures, defining specific rolesets according to the general and idiosyncratic semantic behaviors of predicates in data.

4. Variety of Predicates

Because verbs generally provide the bulk of the event semantics of any given sentence, verbs have been the target of most of the existing two million words of PropBank annotation. Nonetheless, to fully capture event relations, annotations must recognize the potential for their expression in the form of nouns, adjectives and multi-word expressions, such as Light Verb Constructions (LVCs). Within a language and across languages, the same event can be expressed with different syntactic parts of speech, for example:

4. He fears bears.

5. His fear of bears...

6. He is afraid of bears.

Or, for example:

7. He offered to buy the house.

8. His offer to buy the house...

9. He made an offer to buy the house.

---

[3]The English PropBank roleset lexicon can be viewed at: http://verbs.colorado.edu/propbank/framesets-english/.

This has already been acknowledged by FrameNet, but it has recently been necessary to expand PropBank annotations to provide coverage for noun, adjective and complex predicates. PropBank relied heavily on the NomBank frame files in its initial creation of a nominal frame file inventory, but unlike NomBank, PropBank restricts annotations primarily to eventive nouns. WordNet (Fellbaum et al., 1998) and FrameNet have since been referred to frequently in expanding PropBank's inventory of nominal and adjective frame files, and in assessing the derivational relationships between noun, verb and adjective rolesets.

To best leverage the patterns between certain syntactic structures and the expression of a particular event participant, the PropBank annotations are layered on top of the Penn Treebank (Marcus et al., 1994). However, this also makes the semantic annotations of PropBank somewhat constrained by the syntactic representation. Specifically, annotations are limited to the syntactic domain of locality of a particular predicate, and participants expressed outside of that span will not be included in annotations. This also brings up annotation considerations unique to each new predicate type, as each is realized in a very different syntactic environment. These considerations will be discussed in the sections to follow.

4.1. Adjective Predicate Annotation

Previously, when annotating copular verbs followed by an adjectival predicate, PropBank was only able to consider semantics of the verb. For example, in 6 above, '*he is afraid of bears*,' the annotation would capture

10. He-ARG1-TOPIC is-REL [afraid of bears.]-ARG2-COMMENT

There is more going on in this sentence, however. Specifically, a *fearing* event. Previously, given the rather superficial annotation of the verb *be* seen in 10, annotations did not distinguish between this *fearing* event and a *hunger* state in a sentence such as, '*he was hungry*.' This is because, of course, the adjectives in these sentences are more semantically enlightening than the verb. To address this issue, annotation has expanded to include predicate adjectives.

The support verb does play a role in the sentence, so annotation of only the adjective is not enough. Returning to 6 '*he is afraid of bears*,' *he* is syntactically an argument of the verb *be*, rather than an argument of the adjective *afraid*. In order to capture all of the participants of this *fearing* event, PropBank now annotates the support verb and its syntactic domain containing the Experiencer argument, and also the predicating adjective and its syntactic domain containing the Stimulus argument:

11. He-ARG0 is afraid-REL [of bears.]-ARG1

It is important to note that while there is a prepositional phrase *of bears* that is within the domain of the adjective, annotated as an argument of *afraid*, a modifying phrase is not necessary for annotation. For example, *He was hungry* and *He is afraid* would be annotated respectively as:

12. He-ARG0 was hungry-REL

13. He-ARG0 is afraid-REL

Just as with verbs, regardless of whether the predicate adjective is modified, it still carries important semantic information that was lost by only annotating the support verb. Notably, this treatment is more similar to how FrameNet handles adjectives. In FrameNet, adjectives are listed in frames with their participants, or Frame Elements. For example, the adjective *hungry* is listed in the Biological_urge Frame, with an Experiencer of the State (*hungry*) that is parallel to PropBank's Arg0. As PropBank expands the types of predicates annotated and, in turn, the roleset lexicon, a deliberate effort is made to ensure interoperability between PropBank and both FrameNet and VerbNet. Each roleset, or sense, is mapped to both a VerbNet class (where possible, given that VerbNet is restricted to verbs) and a FrameNet frame, and each role is mapped to a VerbNet theta role and a FrameNet frame element. As mentioned previously, this mapping, as well as an annotated corpus illustrating these mappings, is an effort known as SemLink (Palmer, 2009). SemLink is downloadable from this site: https://verbs.colorado.edu/semlink/.

4.2. Noun & Light Verb Construction Annotation

As a recent addition, eventive and stative nouns within and outside of LVCs are now annotated. This initially required the creation of over 2,800 noun rolesets to be added to the PropBank lexicon as the nouns arose for annotation. As mentioned previously, many of these were drawn directly from noun rolesets used in the NomBank project, but were expanded by exploring nouns that WordNet listed with certain types thought to be eventive or stative (e.g. noun.event, noun.act, noun.state).

In a fashion quite similar to adjective annotation, the annotation of complex predicates, such as the LVC *make an offer*, calls for annotation of the syntactic arguments of both the light verb *make* and the noun predicate *offer* (for a full description of PropBank LVC annotation guidelines, see Hwang et al. (2010)). Previously, the importance of the noun in such constructions was ignored, as only verb predicates were annotated. Light verb usages were handled either by simply being lumped in with one of the most dominant, semantically general senses of the verb (e.g. make.01, the creation sense of *make*), or through the designation of a roleset listing specific constructions (e.g. *make a bid*). This practice precluded delving into the deeper semantics of these constructions, largely represented by the noun.

In current practices, annotators identify light verbs and the main noun predicate in an initial verb pass of annotation. In a second pass, annotation is completed for the full span of the complex predicate, using the roleset of the noun. Consider the following example, which uses the *offer* roleset:

Arg0: entity offering
Arg1: commodity, thing offered
Arg2: price
Arg3: benefactive or entity offered to

14. Yesterday-ARGM-TEMPORAL, John-ARG0 made-REL

an offer-REL [to buy the house]-ARG1 [for $350,000]-ARG2

This annotation practice ensures that the complete argument structure of the complex predicate receives annotation, regardless of whether the argument is within the domain of locality of the noun or verb, and that the roles assigned reflect the fact that the bulk of the event semantics stem from the noun.

The annotation of LVCs touches on the complex area of constructions in general, which can be used to extend the relational semantics of predicates in novel ways (see, for example, Goldberg (1995)). The productivity of such constructions allows speakers to creatively express events and states, but this also presents a challenge for all static lexical resources, including the lexicon of PropBank rolesets. Although LVCs are 'semi-productive' (Nickel, 1978), meaning that their productivity is constrained in seemingly idiosyncratic ways, new LVCs do enter the language regularly (e.g. *make a backup*, as in *back up the computer*, in the technical domain; *make the ask*, as in *solicit money*, in the fundraising domain) (see Stevenson et al. (2004) for a discussion of statistical measures of LVC productivity). The previous approach to this lexicon has been to add a new roleset, or sense, whenever a usage that is both syntactically and semantically distinct presents itself in the data. For example, the past approach would have required the addition of a new multi-word roleset for each detected LVC (e.g. *make_offer, take_walk, have_drink, give_sigh, do_investigation*). Such an approach quickly becomes quite cumbersome, and does not address the theoretical problems involved with considering these expressions to be sense extensions of the verb. The current approach therefore makes use of existing noun rolesets instead.

5. Constructions with Adjective Predicates

The problem of productive constructions quickly became intractable in the case of adjective predicate annotations. Specifically, the data was riddled with certain constructions that carry their own semantics and are compatible with essentially any gradable adjective. For example, here are several instances of one of the most common constructions seen in the PropBank corpus of adjective annotations (adjective relations are shown in bold-face):

15. We are too **selfish** to give these programs up

16. I was **stupid** enough to buy all the tripe about what Iraq could become

17. It was a bit too **political** for my liking.

This construction, termed the 'Degree-Consequence' construction for annotators, consists of a gradable adjective that is modified with some kind of degree word (e.g. *enough*, *too*) and an argument indicating what could loosely be thought of as a consequence for the degree to which the state is or isn't true. Notably, this construction is compatible with essentially any gradable adjective, including adjectives that normally would not take any additional arguments beyond the participant characterized by the adjective, like *blue* or *big*:

18. The ball is **blue** enough to be seen in the green grass.

19. The button is too **big** to fit through the hole.

This widespread compatibility, or productivity, evidences the notion that the semantics of the arguments are more likely projected by the construction itself, rather than the predicate adjective. Accommodating such usages forced a choice between creating a roleset that could appropriately label these arguments for every adjective it arose with, or adjusting the guidelines to recognize that constructions themselves can project semantics. Current guidelines adopt the latter approach, with practices being refined as the number and variety of constructions are explored. Currently, arguments that seem to be projected by the construction are marked with a recently developed 'Construction' (CXN) tag. This tag is used in combination with the arguments given in the roleset of the adjective:

20. We-ARG0 are too-ARG-CXN selfish-REL [to give these programs up]-ARG-CXN

Future work will include a second pass to decompose the semantics of these constructions, potentially including the creation of rolesets for each construction. Additionally, verb constructions will be further explored. In this research, the FrameNet Constructicon (Fillmore et al., 2012) will be leveraged to the greatest extent possible, ensuring that PropBank is compatible with the work on constructions found in FrameNet. For example, the FrameNet Constructicon lists a more general Degree construction as well as a subordinate frame (inheriting from the Degree frame), Degree_so. The Degree_so construction is quite similar, but not identical to, the PropBank Degree-Consequence construction: *I am so proud of him; The smell is so terrible you want to throw up*. Ideally, work on constructions in PropBank can help to inform and expand that of FrameNet as well.

6. Predicate Unification & Abstract Meaning Representation

These efforts to address new predicate types not only allow for PropBank to provide greater coverage of event semantics, but also allow for greater interoperability with the related Abstract Meaning Representation (AMR) project. The goal of AMR is quite complementary to PropBank. It aims to create a large-scale semantics bank of simple structures for complete sentences, much in the way the Penn Treebank provided syntactic information. AMR differs from PropBank primarily in the fact that a deliberate attempt is made to abstract away from language-particular syntactic facts, representing instead only concepts and relations in a manner that would ideally allow for meaning-based machine translation. Thus, annotations are done without consideration of the syntactic tree or a particular domain of locality. Additionally, implicit concepts can be included in the meaning representation. For example:

21. Gas could go to $10 a gallon.

In the AMR for this sentence, the implicit concept of 'price' is introduced, for it is actually the price of gas that is rising.

```
(p / possible
  :domain (g / go-01
              :ARG1 (m / monetary-quantity
                         :ARG2-of (p2 / price-01
                                      :ARG1 (g2 / gas
                                                :quant (v / volume-quantity
                                                          :unit (g3 / gallon)
                                                          :quant 1))))
              :ARG4 (m2 / monetary-quantity
                        :unit (d / dollar)
                        :quant 10)))
```

Figure 1: AMR for *Gas could go to $10 a gallon*

The full AMR is given in Figure 1.[4] In this way, AMR has more flexibility in representing the semantics of a sentence where PropBank is somewhat constrained by syntax.

AMR builds upon the foundation of PropBank primarily by using the rolesets that were developed for PropBank. However, AMR makes use of only a single roleset denoting a particular event or state, instead of using a roleset that's tied to a predicate's part of speech, precisely because an effort is made to make a unified representation across different syntactic realizations of the same event. Therefore, where PropBank has three separate rolesets for *fear*-verb, *fear*-noun, and *afraid*-adjective, AMR would generalize all realizations to one roleset, representing the abstract concept of fear. For example, instances 4-6, repeated here for the reader's convenience, would all receive the same simple AMR, illustrated in Figure 2.

22. He fears bears.

23. His fear of bears

24. He is afraid of bears.

```
(f / fear-01
  :ARG0 (h / he)
  :ARG1 (b / bear))
```

Figure 2: AMR for examples 22-24

To better accommodate the use of these rolesets in the AMR project, and to better represent the common concept of a given event across distinct syntactic realizations, PropBank is now unifying rolesets representing the same event or state across different parts of speech. This makes use of a process of 'aliasing,' wherein different lexical items can be aliases of the same concept. Thus, for example, *fear*-noun, *fear*-verb and *afraid*-adjective will be aliases associated with a single fear roleset. Similarly, *offer*-verb, *offer*-noun and the *make_offer*-light verb construction will be aliases of a single *offer* roleset. In this way, PropBank is better equipped to pursue its overarching goal of providing information on event semantics consistently across various syntactic (and morphological) realizations. Previously this goal was pursued on the level of various syntactic realizations of arguments, and now it is pursued on the level of various syntactic realizations of the predicates themselves.

This will provide a more comprehensive and complete view of event relations across a corpus and allow for deeper natural language understanding.

To better illustrate what kinds of changes the unification of existing PropBank frame files will entail, pre- and post-unification examples are given below. Examples 25-26 and 28-31 are pre-unified rolesets that will be subsumed as aliases under their respective unified rolesets, given in 27 and 32.

Event relation: Offer

25. **Predicate**: *offer*-verb
    Roleset id: offer.01 transaction
    Roles: Arg0: entity offering
    Arg1: commodity
    Arg2: price
    Arg3: benefactive or entity offered to
    Example: *He offered to buy the house.*

26. **Predicate**: *offer*-noun
    Roleset id: offer.01 transaction
    Roles: Arg0: entity offering
    Arg1: commodity
    Arg2: price
    Arg3: benefactive or entity offered to
    Example: *His offer to buy the house...*
    He made an offer to buy the house.

27. **UNIFIED ROLESET**
    Predicate aliases: *offer*-verb, *offer*-noun
    Roleset id: offer.01 transaction
    Roles: Arg0: entity offering
    Arg1: commodity
    Arg2: price
    Arg3: benefactive or entity offered to
    Example: *He offered to buy the house.*
    His offer to buy the house..
    He made an offer to buy the house.

```
(o / offer-01
  :ARG0 (h2 / he)
  :ARG1 (b2 / buy-01
             :ARG0 h2
             :ARG1 (h3 / house)))
```

Figure 3: AMR for examples 25-27

[4]For more general information on AMR, see http://amr.isi.edu.

3017

Event relation: Fear

28. **Predicate**: *fear*-verb
    Roleset id: fear.01 fear
    Roles: Arg0: entity afraid
          Arg1: afraid of what?
    Example: *He fears bears.*
    He fears for his life.

29. **Predicate**: *fear*-noun
    Roleset id: fear.01 fear
    Roles: Arg0: entity afraid
          Arg1: afraid of what?
          Arg2: afraid for
    Examples: *His fear of bears...*
    His fear for his life...

30. **Predicate**: *afraid*-adjective
    Roleset id: afraid.01 fear
    Roles: Arg0: entity afraid
          Arg1: afraid of what?
          Arg2: afraid for
    Examples: *He is afraid of bears.*
    He is afraid for his life.

31. **Predicate**: *fearful*-adjective
    Roleset id: fearful.01 afraid
    Roles: Arg0: entity afraid
          Arg1: afraid of what?
    Examples: *He is fearful of bears.*
    He is fearful for his life.

32. **UNIFIED ROLESET**
    Predicate aliases: *fear, afraid, fearful*
    Roleset id: fear.01 fear
    Roles: Arg0: entity afraid
          Arg1: afraid of what?
          Arg2: afraid for
    Examples: *He fears bears.*
    His fear for his life...
    He is afraid of bears.
    He is fearful for his life.

```
(f / fear-01
   :ARG0 (h4 / he)
   :ARG2 (l / life
            :poss h4))
```

Figure 4: AMR for *life* examples 28-32 with Arg2 *life*

Notice that the pre-unified rolesets given in 29 and 30 include an Arg2, while those in 28 and 31 do not. In cases such as this, a decision will be made to either drop or preserve the argument, and any affected past annotation will be adjusted. Arguments will generally be preserved if they contain information that is considered core to the event relation as a whole, and cannot be easily replaced by an modifier tag.

The most apparent difference in FrameNet's handling of these particular event relations compared to PropBank and AMR lies in the way derivationally related lexical units are grouped into rolesets or frames. Whereas Propbank and AMR group together the aliases given in examples 27 and 32, FrameNet splits them between two different frames. *Fear*-noun falls into the 'Fear' frame, *fear*-verb falls into the 'Experiencer_Focus' frame, and afraid-adjective is included in both. *Fearful*-adjective is not represented in FrameNet at all. As a result, because *fear*-noun and *fear*-verb fall into different frames, the example sentences given in 28, 29 and 31 would not be recognized as describing the same eventuality under FrameNet annotation, whereas they would according to PropBank and AMR.

7. Conclusion

PropBank verb annotation has provided over two million words for machine learning. However, verbs alone are not sufficient to express all sentential meaning. Annotation of nouns, adjectives, and multiword expressions has allowed PropBank to capture comparable semantic meanings represented in derivationally different realizations. To provide greater interoperability with the AMR project and to unify the same semantic sense realized in different syntactic environments, PropBank will now make use of aliases to connect these concepts. Additionally, future work will address the innate semantics of generalizable constructions, such as the Degree-Consequence construction found with adjectives.

8. Acknowledgement

9. References

Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Herjakob, Kevin Knight, Philipp Kohen, Martha Palmer, and Nathan Schneider. 2013. Abstract meaning representation for sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*.

Cristiane Fellbaum, Joachim Grabowski, and Shari Landes. 1998. Performance and confidence in a semantic annotation task. In Christiane Fellbaum, editor, *WordNet: An Electronic Database*. The MIT Press.

Charles J. Fillmore, Russell R. Lee-Goldman, and Russell Rhomieux. 2012. The FrameNet Constructicon. In Hans C. Boas and Ivan A. Sag, editors, *Sign-Based Construction Grammar*. CSLI Publications, October.

Ana-Maria Giuglea and Alessandro Moschitti. 2006. Semantic role labelling via FrameNet, VerbNet and PropBank. In *Proceeding of COLING/ACL-2006*, Sydney, Australia.

Adele E. Goldberg. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University Of Chicago Press.

Jena D. Hwang, Archna Bhatia, Claire Bonial, Aous Mansouri, Ashwini Vaidya, Nianwen Xue, and Martha Palmer. 2010. Propbank annotation of multilingual light verb constructions. In *Proceedings of the Fourth Linguistic Annotation Workshop*, pages 82–90, Uppsala, Sweden, July. Association for Computational Linguistics.

Edward Loper, Szu-ting Yi, and Martha Palmer. 2007. Combining lexical resources: Mapping between Prop-Bank and VerbNet. In *Proceedings of the Seventh International Workshop on Computational Semantics (IWCS-7)*.

Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. 1994. Building a large annotated corpus of english: the penn treebank. *Computational Linguistics*, 19.

Paola Merlo and Lonneke van der Plas. 2009. Abstraction and generalisation in semantic role labels: Prop-bank, verbnet or both? In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 288–296, Suntec, Singapore, August.

Adam Meyers, Ruth Reeves, Catherine Macleod, Rachel Szekely, Veronika Zielinska, Brian Young, and Ralph Grishman. 2004. The nombank project: An interim report. In *Proceedings of the NAACL/HLT Workshop on Frontiers in Corpus Annotation*.

Gerhard Nickel. 1978. Complex verbal structures in english. In Dietrich Nehls, editor, *Studies in Descriptive English Grammar*, pages 63–83. Julius Groos.

Martha Palmer, Dan Guildea, and Paul Kingsbury. 2005. The Proposition Bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–105, March.

Martha Palmer. 2009. Semlink: Linking PropBank, VerbNet and FrameNet. In *Proceedings of the Generative Lexicon Conference (GenLex-09)*, Pisa, Italy.

Lei Shi and Rada Mihalcea. 2005. Putting pieces together: Combining framenet, verbnet and wordnet for robust semantic parsing. In *Computational Linguistics and Intelligent Text Processing; Sixth International Conference (CICLing 2005)*, pages 100–111.

Suzanne Stevenson, Afsaneh Fazly, and Ryan North. 2004. Statistical measures of the semi-productivity of light verb constructions. In *Proceedings of the ACL 2004 Workshop on Multiword Expressions: Integrating Processing*.

Szu-ting Yi, Edward Loper, and Martha Palmer. 2007. Can semantic roles generalize across genre. In *Proceedings of HLT/NAACL-2007*.