Deriving Verb-Meaning Clusters from Syntactic Structure

Paul Kingsbury, Karin Kipper
Computer and Information Sciences
University of Pennsylvania

Abstract

This paper presents a methodology for using the argument structure of sentences, as encoded by the PropBank project, to develop clusters of verbs with similar meaning and usage. These clusters can be favorably compared to the classes developed by the VerbNet project. The most interesting cases are those where the clustering methodology suggests new members for VerbNet classes which will then be associated with the semantic predicates for that class.

1 Introduction

There can be no doubt that meaning is dependent upon a great number of factors. The difficulty in higher-level comprehension in Natural Language Processing is due to the identification of these factors as well as the successful implementation of methods for handling them. Early generative models of syntax failed to answer the meaning question because of their assumption that semantics was a natural outcome of structure, thus neglecting the other factors involved. More recent works, such as those found in WordNet (Miller, 1985; Fellbaum, 1998) and its descendents, fall short for the inverse reason: they focus solely on lexical meaning while ignoring structure altogether.

Another factor in the meaning game is the relationship between individual lexical items of the same function. How are all determiners alike, for example, and how are they individually different? Are there subclasses within the class of determiners, and how do these subclasses differ from each other? Human beings are supreme pattern-matchers and assign things to categories almost to a fault. The categories into which semantically loaded items are grouped must therefore be one of the factors leading to higher-level comprehension, and knowledge of what these categories are and how they are built can help solve the greater meaning question.

2 Issue and Previous Work

Much previous work in the domain of classification of lexical items has focused on verbs. This is a natural starting place, since the verb is the hook upon which the rest of a sentence hangs. Verbs also display a higher degree of variation in their semantics than other lexical types. Whereas nouns, for example, all name some kind of 'thing,' verbs can describe an action or a state and involve some number of non-verbal actors in the description. It is this latter quality of verbs which most interests us here, and indeed which has been the focus of most previous work in verb classification. Even the most elementary grammars of English draw a distinction between transitive and intransitive verbs. More advanced work has refined this distinction into ever-larger numbers of classes. For example, the landmark work of Levin (1993) divided over three thousand verbs into 191 classes based partly on shared semantics and partly on shared syntactic alternations. More recently, the VerbNet project at Penn (Kipper et al., 2000) incorporated Levin's verb classification to systematically create verb entries in a lexicon. On a purely semantic note, WordNet (Miller, 1985; Fellbaum, 1998) has classified much of the vocabulary of English, not just verbs, in terms of relationships such as synonymy, hyponymy, and others. Various attempts worldwide have begun focussing on the argument structure of verbs as part of developing dependency grammars. The PropBank project at Penn (Kingsbury and Palmer, 2002) is an example of this process for English; similar projects are

underway for Czech (Hajicova etc), German (Broker 1998), and others. The FrameNet project at Berkeley (Baker et al., 1998) has classified many words in terms of their relation to a relatively small number of core semantic concepts such as 'commerce' and 'judgment'. Various attempts have been made to automatically cluster verbs into semantically meaningful classes, using the Levin class as a gold standard for evaluation (Gildea, 2002; McCarthy, 2000; Merlo and Stevenson, 2001; Schulte im Walde, 2000). In the next two sections, we provide background on VerbNet and PropBank which play central roles in the cluster methodology presented here.

2.1   VerbNet

VerbNet is a verb lexicon with syntactic and semantic information for English verbs, referring to Levin verb classes (Levin, 1993) for systematic construction of lexical entries. This lexicon exploits the systematic link between syntax and semantics that motivates these classes, and thus provides a clear and regular association between syntactic and semantic properties of verbs and verb classes (Kipper et al., 2000; Dang et al., 2000). Each class in the hierarchy is composed of a set of members (linked to their WordNet synsets) and a set of syntactic frames and semantic information for each frame. Currently, VerbNet has over 4,000 verb senses described (3,004 lemmas) within 191 first level classes.

VerbNet has a hierarchical structure, with the first level classes constituted by the original Levin classes. In order to ensure that each class is coherent, so that all its members share a common set of thematic roles, syntactic frames and semantic predicates, some restructuring of the classes was required. This reorganization, which was facilitated by the use of intersective Levin classes (Dang et al., 1998), refined the classes to account for semantic and syntactic differences within a class. A child subclass inherits all the information from its parent class, and adds information to it, which can be in terms of imposing further restrictions on the roles, or adding syntactic frames or semantic predicates to the subclass.

The hierarchical organization of VerbNet is illustrated in Figure 1. The *Transfer of a Message* verb class is subdivided into three levels. At the top level are thematic roles, syntactic frames and semantic predicates shared by all members of the class. In this particular case, there is a transitive frame with the Topic (message) as the direct object (Agent Verb Topic), as in "John explained trigonometry", and a frame for Topic and Recipient (Agent Verb Topic to Recipient), as in "John taught math to Mary". Both syntactic frames have semantic predicates expressing

the transfer of information event, but in the first case the Recipient is underspecified. Some of the verbs in this class are able to participate in other syntactic frames as well. Verbs at the second level can take the ditransitive frame (Agent Verb Recipient Topic) in addition to the frames and predicates inherited from the parent class.

VerbNet uses a flat semantic representation in which the semantics of each syntactic frame is captured by a conjunction of predicates [1], such as *motion, contact, transfer_info*, which can be negated or not. These predicates can take arguments over the verb complements, as well as over implicit existentially quantified event variables.

Each semantic predicate in VerbNet also include a time function specifying whether the predicate is true in the preparatory ($during(E)$), culmination ($end(E)$), or consequent ($result(E)$) stage of an event, in a tripartite event structure is similar to that of Moens and Steedman (1988), which allows us to express the semantics of classes of verbs like change of state verbs whose description requires reference to a complex event structure.

2.2   PropBank

In a different vein, the PropBank project (Kingsbury and Palmer, 2002) has endeavoured to describe all the most frequent verbs of English in terms of their argument structure. This project has three major differences from previous works. First, the description of each verb is accompanied by a rich set of examples drawn from real language, in this case the Wall Street Journal sections of the Penn Treebank (Marcus, 1994). Furthermore, the descriptions are based on the usages in the corpus, rather than a possible situation where the corpus was mined for sentences fitting preconceived patterns. This results in many instances which are perfectly well-formed but unexpected. The best example of this is an odd usage of the verb *add*: *The Nasdaq composite index added 1.01 to 456.6 on paltry volume.*, where the context makes it clear that *add* is being used as a synonym for *rise*. Second, argument structure allows for a richer set of descriptions than merely 'transitive', 'unaccusative', and so forth, since any individual verb is allowed to have anywhere between zero and six arguments. Third and perhaps most importantly, the PropBank descriptions make explicit mention of the different senses of verbs. This is crucial because different senses can have different argument structures or different syntactic alternations, a detail which is often glossed over in other resources. Thus, while (1)

---

[1] Presently there are 64 distinct predicates described.

Transfer of a Message - level 1 class

⟨⟨MEMBERS ⟩⟩		[⟨cite, wn1⟩, ⟨demonstrate, wn1⟩, ⟨explain, wn1⟩, . . .]
⟨⟨THEMATIC ROLES ⟩⟩		Agent(A), Recipient(R), Topic(T)
⟨⟨SELECT RESTRICTIONS ⟩⟩		Agent[+animate], Recipient[+animate], Topic[+message]

⟨⟨FRAMES and PREDICATES ⟩⟩

Transitive with Topic	A V T	transfer_info(during(E),A,?,T)
Topic and Recipient	A V T to R	transfer_info(during(E),A,R,T)

Transfer of a Message - level 2 class

| ⟨⟨PARENT ⟩⟩ | | Transfer of a Message - level 1 |
| ⟨⟨MEMBERS ⟩⟩ | | [⟨dictate, wn2⟩, ⟨quote, wn1⟩, ⟨read, wn3⟩] |

⟨⟨FRAMES and PREDICATES ⟩⟩

Ditransitive	A V R T	transfer_info(during(E),A,R,T)

Figure 1: Example entries for the *Transfer of a Message - levels 1* and *2* classes

and (3) share the same argument structure, (2) and (4) do not, because of the ungrammaticality of (4):

1 Congress passed the bill.

2 John passed the butter.

3 The bill passed.

4 *The butter passed.

Under the PropBank scheme, the similarity between *pass* (make bill into law) and *pass* (move from one place to another) is regarded an almost-accidental byproduct of centuries of semantic drift. The combination of accounting for all the usages in a large corpus and separating verb senses results in a very large database: 4358 distinct senses spread across 3183 lexical items, with a total of over 70 000 unique tokens. PropBank annotations are recorded as theory-neutrally as possible. Crucially, the nomenclature of traditional thematic roles is not used; instead, simple labels of Arg0, Arg1 and so forth are used. Each ArgX is mapped to a specific role on a per-verb basis, such as the roleset for *give*:

Arg0: giver
Arg1: gift
Arg2: given-to
ex: [$_{Arg0}$ The executives] gave [$_{Arg2}$ the chefs] [$_{Arg1}$ a standing ovation].

Such a schema has its advantages and its drawbacks. On the negative side, it makes it difficult to draw comparisons between various arguments with the same number. The meaning of Arg2 for one verb might bear no relation to the meaning of Arg2 for some other verb, for example. On the plus side, not using traditional thematic roles both frees the annotators from the necessity of distinguishing between Theme and Patient, for example, while also allowing for roles which do not have corresponding thematic roles. For example, PropBank has a host of verbs relating to the judicial process, such as *jail, sentence, imprison*, and so forth. All of these take four roles;

Arg0: lawgiver
Arg1: guilty party
Arg2: term, length of jail time
Arg3: crime
ex: Dallas District Judge Jack Hampton had sparked calls for a judicial inquiry with his remarks to the press last December, two weeks after [$_{Arg0}$ *trace=Hampton*] sentencing [$_{Arg1}$ an 18-year-old defendant] to [$_{Arg2}$ 30 years in state prison] for [$_{Arg3}$ killing two homosexual men in a city park.]

While Args 0 and 1 can be thought of as Agent and Patient, traditional thematic role terminology fails to accommodate Args 2 and 3, for all that they are central to the meaning of the verbs.

3 Methodology

The PropBank annotations thus provide a rich set of actual usages which are easily divided into senses and quantified. That is, for any verb (sense) it is possible to collect all the usages, divide them into their various syntactic realizations, and enumerate these realizations. The realizations themselves are

rendered at a very basic level, encoding only the linear order of arguments. The verb itself is rendered as *rel* for 'relation'. Thus a verb such as *repeal* can appear in either a simple transitive or passive:

> [$_{Arg0}$ West Germany] will repeal [$_{Arg1}$ the unpopular turnover tax on security transactions]. (wsj_0302)
> → Arg0 rel Arg1

> ...the tax will be officially repealed [$_{Arg1}$ *trace=the tax*] before the end of the current parliamentary term... (wsj_0302)
> → rel Arg1

As seen in the second example, adjuncts such as the temporal *before the end of the current parliamentary term* are omitted. While the bulk of annotations fall into basic categories such as Arg0 rel Arg1 (basic transitive), Arg0 rel Arg1 Arg2 (ditransitive) and so forth, the nature of the data and the annotation process itself means that there is a large residue of irregular syntactic realizations. As an example of the former, consider the following sentence, where the benefactive argument has been topicalized to the beginning:

> For [$_{Arg4}$ Mr. Sherwin], [$_{Arg0}$ a conviction] could carry [$_{Arg1}$ penalties of five years in prison and a \$250,000 fine on each count]. (wsj_1331) → Arg4 Arg0 rel Arg1

Such a construction is certainly a natural part of English, yet it occurs very rarely—only twice in all of Propbank (so far). The annotation process itself also introduces many odd, low-frequency syntactic realizations which should be regarded as errors. For example, occasionally annotators include one argument label twice:

> ...the board also was informed...of [$_{Arg1}$ interest] expressed by [$_{Arg0}$ buy-out funds including Kohlberg Kravis Roberts & Co. and Forstmann Little & Co.], as well as by [$_{Arg0}$ Robert Bass, Morgan Stanley's buy-out fund, and Pan Am Corp]. (wsj_2104) → Arg1 rel Arg0 Arg0

In cases of such conjoined constructions the annotation guidelines specify that the sentence should be regarded as two overlapping sentences, with two overlapping argument structures. Thus, for the sentence above, there should properly be two instances of 'Arg1 rel Arg0', each with the same Arg1 and rel

but with different Arg0's. The annotator of this sentence clearly forgot this specification. Fortunately, PropBank is proceeding through a double blind annotation followed by an adjudication phase. The adjudication will catch and correct structures such as the above. For the time being, however, such errors are retained, largely for want of an efficient and accurate filtering mechanism.

Other filtering on the data is performed, however. Since the TreeBank corpus is subject to the same Zipfian distribution[2] of lexical items as any natural-language database, many verbs are only poorly attested. A largely arbitrary decision was made to eliminate all verbs which had fewer than 10 attestations. Finally, again due to the ongoing nature of the PropBank work, not all polysemous verbs have been disambiguated. These verbs were also deleted from the subsequent analysis. All the same, the resulting dataset provides 921 verbs with 200 distinct syntactic realizations of varying frequencies. Each realization was recorded as a proportion of the number of attestations of that syntactic pattern to the total number of attestations of that verb. Expressing these as proportions rather than as raw counts allows for better comparisons among verbs which themselves vary widely in frequency. The proportions themselves are then run through a standard clustering algorithm implemented in R (http://www.R-project.org) which classifies each verb by the difference between its attested proportions and the proportions attested by every other verb. Verbs with little difference, meaning they attest nearly the same syntactic patterns in nearly the same proportions, are judged to be very similar and are likely to be grouped into the same class. A varying number of "centroids" or prototypes for each class can be established. Each verb is then classified by its similarity to the centroids, resulting in a number of classes equal to the number of centroids. The resulting classes are compared to the existing VerbNet classes. This is at best a noisy measure of accuracy, since many of the Prop-Banked verbs (even fairly richly-attested ones) are not classified in VerbNet, and many verbs which are classified in VerbNet are not present in PropBank in sufficient numbers to undergo the clustering analysis. These factors are constant across all possible number of centroids/classes, however, so at least the amount of noise in the assessment remains constant as well.

---

[2]The probability of occurrence of words or other items starts high and tapers off. Thus, a few occur very often while many others occur rarely.

4 Results

As an initial assessment of the worth of the methodology, consider the case with only three clusters. The clusters can be viewed graphically in Figure 2.

The crosses are the centroids. Such a display as this is naturally a gross simplification of the actual 200-dimensional space the clusters are mapped in, but it's the best that can be done on paper. If we look at (a sample of) the membership of each of the three clusters, we see that the three groups clearly fall into the sets of transitive, unaccusative, and ditransitive verbs:[3]

Class 1	Class 2	Class 3
waste	repay	drift
ask.01	invest	die
overcome	appoint	crumble
feel.01	invite	triple
make.05	assign	float
influence	recruit	disappear

For higher numbers of clusters a more formal method of measuring accuracy is required. Since we are comparing sets, the similarity of two sets A and B can be computed with the following:

$$similarity(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

In other words, the similarity of two sets can be measured as the number of elements shared by the two sets divided by the total number of unique elements in the two sets. Thus, two identical sets ({a, b} {b, a}) would have a similarity score of 1, because the intersection contains two elements, as does the union. Two sets a, b and a, b, ..., z would have a very low similarity score, since the intersection contains two elements, but that is divided by the union with 26 elements. Thus the similarity score is dependent both on the number of matches and the number of spurious elements present. For example, the three clusters described above most closely align to the VerbNet classes 13.5.1 (*Get* class), 45.6 (*Calibratable_change_state* class), and 48.1.1 (*Appear class*), respectively, with 4, 4, and 7 matching elements. Each of these clusters is huge, however (521, 335, and 65 verbs, respectively), and the target sets are also large as VerbNet clusters go (31, 23, and 27 verbs, respectively). These factors conspire to make the overall similarity scores quite low (4/(521+31),

---

[3]ask.01: ask a question, as opposed to ask a favor; feel.01: feel emotion; make.05: make money, as in "Paul doesn't make nearly enough money; triple: while in most English this would be considered a transitive verb, the common usage in the Wall Street Journal is actually unacusative, as in "Paul's net worth tripled when he found a $5 bill in the dryer'."

etc). This demonstrates how a three-cluster analysis, while interesting for describing a gross syntactic grouping, does not go very far in describing any finer-grained meaning contrasts.

To get at the classification with real implications for meaning we clearly need to use more, smaller clusters. How many more, and how small? Since the original set of verbs under analysis comes from 150 VerbNet classes, we can set an upper bound on the number of clusters at 150. Then we iteratively derive clusters for each number between 3 and 150 and find the similarity scores for each cluster in each of those sets. To facilitate comparison between the various sets of clusters, we calculate the average similarity across all the classes. This is necessary because, as the number of clusters rises, the size of any individual cluster is likely to fall, thus bringing the denominator of the similarity equation down. Taken across all the clusters, then, the similarity scores are bound to rise regardless of improvements in the matching between the VerbNet clusters and the automatically-derived ones. The outcome of this procedure can be seen in Figure 3.

Interesting peaks appear for analyses with 14 clusters and around 89 clusters before the long tailing off with more than 90 clusters. In addition, there is a local maximum around 32 clusters. The analyses with 14 clusters still uses very large clusters (average of 65 verbs/cluster), so that number is less useful. The smallest cluster developed therein, however, consists only of 'remain' and 'stay.01',[4] and that cluster correctly maps to VerbNet class 47.1-1 (*Exist* class), which contains verbs of staying in place.

By the time around 90 clusters are used the derived classes themselves tend to be very small, with 34 clusters containing two or fewer verbs. In contrast, a few clusters continue to have very large membership, such as two groups (arbitrarily labeled '3' and '38') which contain 145 and 146 verbs, respectively, and which correspond roughly to VerbNet classes 29.5 (*Conjecture* class) and 29.2 (*Characterize* class). While 29.2 is one of the largest classes in VerbNet, precisely why this class should seem to attract vagrant verbs remains to be seen. More interesting, however, are the times when the clustering process establishes multiple clusters which are all identified with the same VerbNet class – in effect, suggesting subclasses where they have not previously appeared. This trend appears even with fairly low numbers of clusters, and it tends to be extremely consistent as to which VerbNet clusters are split. For example, class 45.4 (*Other_change_state*), containing the 'miscella-
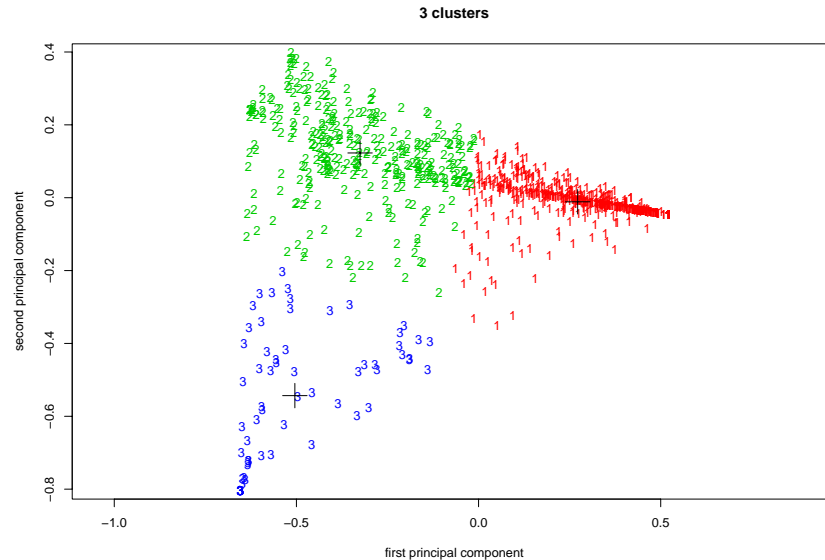
---

[4]Sense:remain.

Figure 2: Clusters for transitive, unaccusative, and ditransitive

neous causatives,' is first split into two subclasses even when only 9 clusters are in play, indicating that there is more variation in the syntactic patterns within this single class than there is between almost any other arbitrarily-selected pair of classes. It is hardly surprising that a collection labeled 'miscellaneous' should contain such a wide variety of syntactic patterns. What is surprising, and encouraging, is the speed at which this methodology identifies this miscellany and moves to correct it.

In one example of how previously-unclassified verbs may be added to a VerbNet class, consider cluster 20 within the 90-cluster analysis. This grouping is optimally identified with VerbNet class 36.3 (*Meet* class), verbs of combative meetings. It contains, unsurprisingly, verbs such as *fight* and *consult*. The clustering analysis adds, in addition, verbs such as *pull.02* (phrasal: *pull out*) and *withdraw*. The former, like all phrasal verbs, is neglected by VerbNet, while the latter is considered to be part of the 'remove' classes (10.1 and 10.5). Yet it might be just as natural to think of *withdraw* in the sense of exiting a meeting or an engagement rather than removing something from somewhere. For example, *"It was just another one of the risk factors" that led to the company's decision to withdraw from the bidding, he added* (wsj_0013). Another verb placed into this cluster is *hedge*. Impressionistically, this might appear to be restricted to "hedge one's bets" and thus a poor match for this fairly combative class. Upon examination of the actual usages, however, it becomes clear that most of the time *hedge* is being used to mean

protect (from), as in *But some investors might prefer a simpler strategy than hedging their individual holdings* (wsj_1962) or *Thus, buying puts after a big market slide can be an expensive way to hedge against risk* (wsj_2415). Seen this way, *hedge* is a plausible member of this class, and since it is hitherto untreated by VerbNet, the clustering analysis provides a valuable suggestion for an unexpected meaning. It also avoids erroneously attributing a change of location predicate to the sentence.

For another example, consider cluster 25 of the same 90-cluster analysis. The verbs in this group contain verbs of directed motion such as *put.01* and *funnel*, and is identified with VerbNet class 9.3 (*Funnel* class). The VerbNet description for this class specifies three arguments: Agent, Theme, and Destination. The Destination role is further specified as being restricted to those cases where the Theme and Destination end up in close contact. Thus, this class includes usages such as *put the toy in the box* but not *put the money towards a new car*. Another verb which is placed into this class is *cast.01*, which is restricted to a small number of near-idioms such as *cast a pall (over)* and *cast doubt (on)*. Because of this restriction, it has not previously been handled by VerbNet. The clustering successfully includes this in the same class as *put* and *funnel*, even including the selectional restriction on the types of Destination allowed, for only *over* and *on* will be found with this sense of *cast*. This would also assign the correct semantic predicates to *cast*.
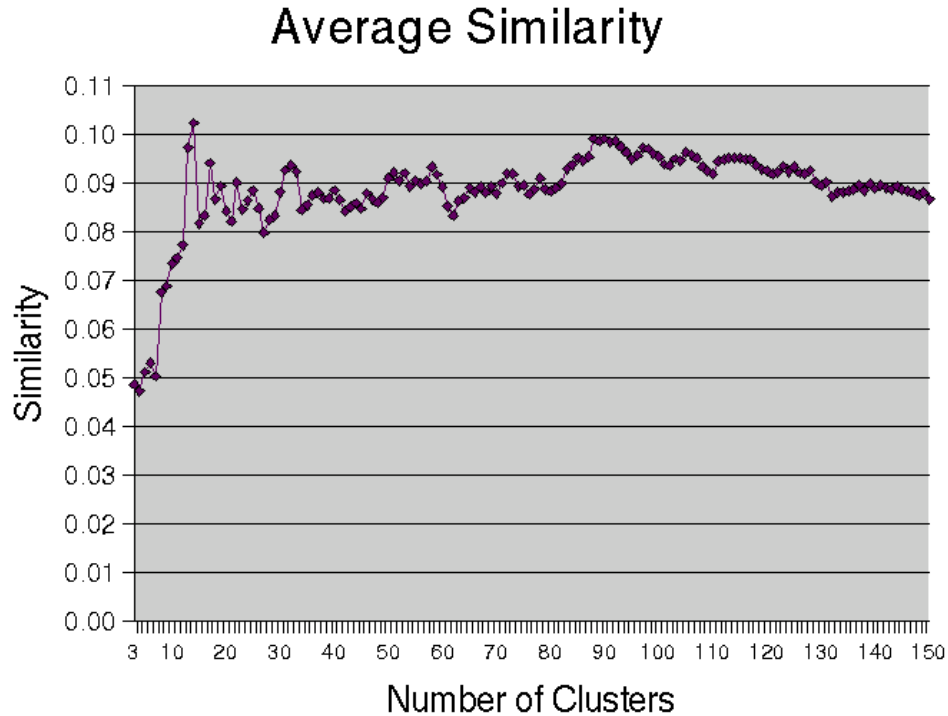
Average Similarity

Figure 3: Outcome of clustering procedure

5   Conclusion and Future Directions

Clustering of syntactic patterns can quickly and automatically provide a first approximation of the groupings into which meaning-bearing items fall. The methodology easily captures low-level generalizations and points out shortcomings in the 'gold standard' verb classification against which it was judged. Nevertheless, it is only a first pass at transforming syntactic structure into semantics. Additional information is available to extend the paradigm. For example, the syntactic patterns which were the subject of the clustering algorithm were nothing more than the linear order of the arguments of the verb. Further refinement on the syntactic structure would be possible, at the expense of multiplying the patterns. For instance, in the preceding all prepositions were stripped from the argument structure, even though PropBank includes the preposition as part of the argument label rather than as part of the argument itself. Also, it would be possible to label the arguments with thematic roles instead of the generic labels used by PropBank, when possible, thus providing a measure of semantics mixed in with the syntax. Kipper et al. (submitted, 2003) describes the process of mapping PropBank argument labels onto VerbNet thematic roles. Cur-

rently, over 64% of the PropBank argument labels have been mapped automatically, and that is without recourse to the sense-disambiguated data. This suggests that the two resources are highly compatible and that the combined data could easily be used in the clustering methodology. A larger measure of meaning could be added if some notion of the meaning of the verbs could be included in the clustering as well. For example, WordNet synonym sets, especially as encoded in VerbNet, could be used as the initial state of the clusters, which then could be modified by the dis/similarity of the syntactic patterns attested by those verbs. Such refinements to the process would in themselves be steps towards integrating the many factors that contribute to the understanding of meaning.

References

Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of the 17th International Conference on Computational Linguistics (COLING/ACL-98)*, pages 86–90, Montreal. ACL.

Hoa Trang Dang, Karin Kipper, Martha Palmer, and Joseph Rosenzweig. 1998. Investigating regular

sense extensions based on intersective levin classes. In *Proceedings of ACL98*, Montreal, Canada, August.

Hoa Trang Dang, Karin Kipper, and Martha Palmer. 2000. Integrating compositional semantics into a verb lexicon. In *Proceedings of the Eighteenth International Conference on Computational Linguistics (COLING-2000)*, Saarbrücken, Germany, July-August.

Christiane Fellbaum, editor. 1998. *WordNet: An Eletronic Lexical Database*. Language, Speech and Communications. MIT Press, Cambridge, Massachusetts.

Daniel Gildea. 2002. Probabilistic models of verb-argument structure. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING-02)*, pages 308–314, Taipei.

Paul Kingsbury and Martha Palmer. 2002. From treebank to propbank. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002)*, Las Palmas, Canary Islands, Spain.

Karin Kipper, Hoa Trang Dang, and Martha Palmer. 2000. Class-based construction of a verb lexicon. In *Proceedings of the Seventh National Conference on Artificial Intelligence (AAAI-2000)*, Austin, TX, July-August.

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

Mitch Marcus. 1994. The penn treebank: A revised corpus design for extracting predicate-argument structure. In *Proceedings of the ARPA Human Language Technology Workshop*, Princeton, NJ, March.

Diana McCarthy. 2000. Using semantic preferences to identify verbal participation in role switching alternations. In *Proceedings of the 1st Annual Meeting of the North American Chapter of the ACL (NAACL)*, pages 256–263, Seattle, Washington.

Paola Merlo and Suzanne Stevenson. 2001. Automatic verb classification based on statistical distribution of argum ent structure. *Computational Linguistics*, 27(3), September.

George Miller. 1985. Wordnet: A dictionary browser. In *Proceedings of the First International Conference on Information in Data*, Waterloo, Ontario.

M. Moens and M. Steedman. 1988. Temporal Ontology and Temporal Reference. *Computational Linguistics*, 14:15–38.

Sabine Schulte im Walde. 2000. Clustering verbs semantically according to their alternation behaviour. In *In Proceedings of the 18th International Conference on Computational Linguistics (COLING-00)*, pages 747–753, Saarbrücken, Germany.