

Classifying Particle Semantics in English Verb-Particle Constructions

Paul Cook

Department of Computer Science
University of Toronto
Toronto, ON M5S 3G4
Canada
pcook@cs.toronto.edu

Suzanne Stevenson

Department of Computer Science
University of Toronto
Toronto, ON M5S 3G4
Canada
suzanne@cs.toronto.edu

Abstract

Previous computational work on learning the semantic properties of verb-particle constructions (VPCs) has focused on their compositionality, and has left unaddressed the issue of *which* meaning of the component words is being used in a given VPC. We develop a feature space for use in classification of the sense contributed by the particle in a VPC, and test this on VPCs using the particle *up*. The features that capture linguistic properties of VPCs that are relevant to the semantics of the particle outperform linguistically uninformed word co-occurrence features in our experiments on unseen test VPCs.

1 Introduction

A challenge in learning the semantics of multiword expressions (MWEs) is their varying degrees of *compositionality*—the contribution of each component word to the overall semantics of the expression. MWEs fall on a range from fully compositional (i.e., each component contributes its meaning, as in *frying pan*) to non-compositional or idiomatic (as in *hit the roof*). Because of this variation, researchers have explored automatic methods for learning whether, or the degree to which, an MWE is compositional (e.g., Lin, 1999; Bannard et al., 2003; McCarthy et al., 2003; Fazly et al., 2005).

However, such work leaves unaddressed the basic issue of which of the possible meanings of a component word is contributed when the MWE is (at least partly) compositional. Words are notoriously ambiguous, so that even if it can be determined that an MWE is compositional, its meaning

is still unknown, since the actual semantic contribution of the components is yet to be determined. We address this problem in the domain of verb-particle constructions (VPCs) in English, a rich source of MWEs.

VPCs combine a verb with any of a finite set of particles, as in *jump up*, *figure out*, or *give in*. Particles such as *up*, *out*, or *in*, with their literal meaning based in physical spatial relations, show a variety of metaphorical and aspectual meaning extensions, as exemplified here for the particle *up*:

- (1a) The sun just came up. [vertical spatial movement]
- (1b) She walked up to him. [movement toward a goal]
- (1c) Drink up your juice! [completion]
- (1d) He curled up into a ball. [reflexive movement]

Cognitive linguistic analysis, as in Lindner (1981), can provide the basis for elaborating this type of semantic variation.

Given such a sense inventory for a particle, our goal is to automatically determine its meaning when used with a given verb in a VPC. We classify VPCs according to their particle sense, using statistical features that capture the semantic and syntactic properties of verbs and particles. We contrast these with simple word co-occurrence features, which are often used to indicate the semantics of a target word. In our experiments, we focus on VPCs using the particle *up* because it is highly frequent and has a wide range of meanings. However, it is worth emphasizing that our feature space draws on general properties of VPCs, and is not specific to this particle.

A VPC may be ambiguous, with its particle occurring in more than one sense; in contrast to (1a), *come up* may use *up* in a goal-oriented sense as in

The deadline is coming up. While our long-term goal is token classification (disambiguation) of a VPC in context, following other work on VPCs (e.g., Bannard et al., 2003; McCarthy et al., 2003), we begin here with the task of type classification. Given our use of features which capture the statistical behaviour relevant to a VPC across a corpus, we assume that the outcome of type classification yields the predominant sense of the particle in the VPC. Predominant sense identification is a useful component of sense disambiguation of word tokens (McCarthy et al., 2004), and we presume our VPC type classification work will form the basis for later token disambiguation.

Section 2 continues the paper with a discussion of the features we developed for particle sense classification. Section 3 first presents some brief cognitive linguistic background, followed by the sense classes of *up* used in our experiments. Sections 4 and 5 discuss our experimental set-up and results, Section 6 related work, and Section 7 our conclusions.

2 Features Used in Classification

The following subsections describe the two sets of features we investigated. The linguistic features are motivated by specific semantic and syntactic properties of verbs and VPCs, while the word co-occurrence features are more general.

2.1 Linguistically Motivated Features

2.1.1 Slot Features

We hypothesize that the semantic contribution of a particle when combined with a given verb is related to the semantics of that verb. That is, the particle contributes the same meaning when combining with any of a semantic class of verbs.¹ For example, the VPCs *drink up*, *eat up* and *gobble up* all draw on the completion sense of *up*; the VPCs *puff out*, *spread out* and *stretch out* all draw on the extension sense of *out*. The prevalence of these patterns suggests that features which have been shown to be effective for the semantic classification of verbs may be useful for our task.

We adopt simple syntactic “slot” features which have been successfully used in automatic semantic classification of verbs (Joanis and Stevenson,

¹Villavicencio (2005) observes that verbs from a semantic class will form VPCs with similar sets of particles. Here we are hypothesizing further that VPCs formed from verbs of a semantic class draw on the same meaning of the given particle.

2003). The features are motivated by the fact that semantic properties of a verb are reflected in the syntactic expression of the participants in the event the verb describes. The slot features encode the relative frequencies of the syntactic slots—subject, direct and indirect object, object of a preposition—that the arguments and adjuncts of a verb appear in. We calculate the slot features over three contexts: all uses of a verb; all uses of the verb in a VPC with the target particle (*up* in our experiments); all uses of the verb in a VPC with any of a set of high frequency particles (to capture its semantics when used in VPCs in general).

2.1.2 Particle Features

Two types of features are motivated by properties specific to the semantics and syntax of particles and VPCs. First, Wurmbrand (2000) notes that compositional particle verbs in German (a somewhat related phenomenon to English VPCs) allow the replacement of their particle with semantically similar particles. We extend this idea, hypothesizing that when a verb combines with a particle such as *up* in a particular sense, the pattern of usage of that verb in VPCs using all other particles may be indicative of the sense of the target particle (in this case *up*) when combined with that verb. To reflect this observation, we count the relative frequency of any occurrence of the verb used in a VPC with each of a set of high frequency particles.

Second, one of the striking syntactic properties of VPCs is that they can often occur in either the joined configuration (2a) or the split configuration (2b):

(2a) Drink up your milk! He walked out quickly.

(2b) Drink your milk up! He walked quickly out.

Bolinger (1971) notes that the joined construction may be more favoured when the sense of the particle is not literal. To encode this, we calculate the relative frequency of the verb co-occurring with the particle *up* with each of 0–5 words between the verb and *up*, reflecting varying degrees of verb-particle separation.

2.2 Word Co-occurrence Features

We also explore the use of general context features, in the form of word co-occurrence frequency vectors, which have been used in numerous approaches to determining the semantics of a target

word. Note, however, that unlike the task of word sense disambiguation, which examines the context of a target word token to be disambiguated, here we are looking at aggregate contexts across all instances of a target VPC, in order to perform type classification.

We adopt very simple word co-occurrence features (WCFs), calculated as the frequency of any (non-stoplist) word within a certain window left and right of the target. We noted above that the target particle semantics is related both to the semantics of the verb it co-occurs with, and to the occurrence of the verb across VPCs with different particles. Thus we not only calculate the WCFs of the target VPC (a given verb used with the particle *up*), but also the WCFs of the verb itself, and the verb used in a VPC with any of the high frequency particles. These WCFs give us a very general means for determining semantics, whose performance we can contrast with our linguistic features.

3 Particle Semantics and Sense Classes

We give some brief background on cognitive grammar and its relation to particle semantics, and then turn to the semantic analysis of *up* that we draw on as the basis for the sense classes in our experiments.

3.1 Cognitive Grammar and Schemas

Some linguistic studies consider many VPCs to be idiomatic, but do not give a detailed account of the semantic similarities between them (Bolinger, 1971; Fraser, 1976; Jackendoff, 2002). In contrast, work in cognitive linguistics has claimed that many so-called idiomatic expressions draw on the compositional contribution of (at least some of) their components (Lindner, 1981; Morgan, 1997; Hampe, 2000). In cognitive grammar (Langacker, 1987), non-spatial concepts are represented as spatial relations. Key terms from this framework are:

Trajector (TR) The object which is conceptually foregrounded.

Landmark (LM) The object against which the TR is foregrounded.

Schema An abstract conceptualization of an experience. Here we focus on schemas depicting a TR, LM and their relationship in both the initial configuration and the final configuration communicated by some expression.

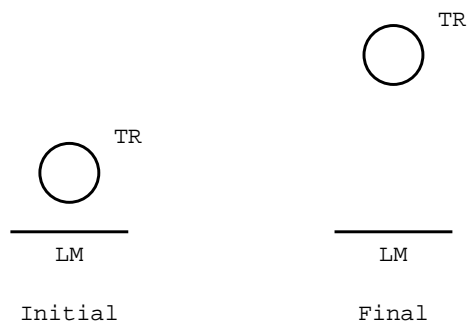


Figure 1: Schema for Vertical *up*.

The semantic contribution of a particle in a VPC corresponds to a schema. For example, in sentence (3), the TR is the balloon and the LM is the ground the balloon is moving away from.

(3) The balloon floated up.

The schema describing the semantic contribution of the particle in the above sentence is shown in Figure 1, which illustrates the relationship between the TR and LM in the initial and final configurations.

3.2 The Senses of *up*

Lindner (1981) identifies a set of schemas for each of the particles *up* and *out*, and groups VPCs according to which schema is contributed by their particle. Here we describe the four senses of *up* identified by Lindner.

3.2.1 Vertical *up* (Vert-up)

In this schema (shown above in Figure 1), the TR moves away from the LM in the direction of increase along a vertically oriented axis. This includes prototypical spatial upward movement such as that in sentence (3), as well as upward movement along an abstract vertical axis as in sentence (4).

(4) The price of gas jumped up.

In Lindner's analysis, this sense also includes extensions of upward movement where a vertical path or posture is still salient. Note that in some of these senses, the notion of verticality is metaphorical; the contribution of such senses to a VPC may not be considered compositional in a traditional analysis. Some of the most common sense extensions are given below, with a brief justification as to why verticality is still salient.

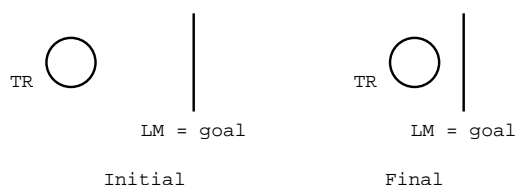


Figure 2: Schema for Goal-Oriented *up*.

Up as a path into perceptual field. Spatially high objects are generally easier to perceive. Examples: *show up*, *spring up*, *whip up*.

Up as a path into mental field. Here *up* encodes a path for mental as opposed to physical objects. Examples: *dream up*, *dredge up*, *think up*.

Up as a path into a state of activity. Activity is prototypically associated with an erect posture. Examples: *get up*, *set up*, *start up*.

3.2.2 Goal-Oriented *up* (Goal-*up*)

Here the TR approaches a goal LM; movement is not necessarily vertical (see Figure 2). Prototypical examples are *walk up* and *march up*. This category also includes extensions into the social domain (*kiss up* and *suck up*), as well as extensions into the domain of time (*come up* and *move up*), as in:

- (5a) The intern kissed up to his boss.
 (5b) The deadline is coming up quickly.

3.2.3 Completive *up* (Cmpl-*up*)

Cmpl-*up* is a sub-sense of Goal-*up* in which the goal represents an action being done to completion. This sense shares its schema with Goal-*up* (Figure 2), but it is considered as a separate sense since it corresponds to uses of *up* as an aspectual marker. Examples of Cmpl-*up* are: *clean up*, *drink up*, *eat up*, *finish up* and *study up*.

3.2.4 Reflexive *up* (Refl-*up*)

Reflexive *up* is a sub-sense of Goal-*up* in which the sub-parts of the TR are approaching each other. The schema for Refl-*up* is shown in Figure 3; it is unique in that the TR and LM are the same object. Examples of Refl-*up* are: *bottle up*, *connect up*, *couple up*, *curl up* and *roll up*.



Figure 3: Schema for Reflexive *up*.

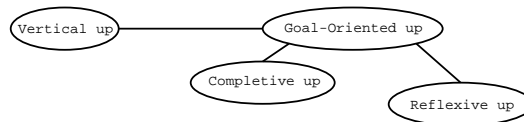


Figure 4: Simplified schematic network for *up*.

3.3 The Sense Classes for Our Study

Adopting a cognitive linguistic perspective, we assume that all uses of a particle make some compositional contribution of meaning to a VPC. In this work, we classify target VPCs according to which of the above senses of *up* is contributed to the expression. For example, the expressions *jump up* and *pick up* are designated as being in the class Vert-*up* since *up* in these VPCs has the vertical sense, while *clean up* and *drink up* are designated as being in the class Cmpl-*up* since *up* here has the completive sense. The relations among the senses of *up* can be shown in a “schematic network” (Langacker, 1987). Figure 4 shows a simplification of such a network in which we connect more similar senses with shorter edges. This type of analysis allows us to alter the granularity of our classification in a linguistically motivated fashion by combining closely related senses. Thus we can explore the effect of different sense granularities on classification.

4 Materials and Methods

4.1 Experimental Expressions

We created a list of English VPCs using *up*, based on a list of VPCs made available by McIntyre (2001) and a list of VPCs compiled by two human judges. The judges then filtered this list to include only VPCs which they both agreed were valid, resulting in a final list of 389 VPCs. From this list, training, verification and test sets of sixty VPCs each are randomly selected. Note that the expense of manually annotating the data (as described below) prevents us from using larger datasets in this initial investigation. The experimental sets are

chosen such that each includes the same proportion of verbs across three frequency bands, so that the sets do not differ in frequency distribution of the verbs. (We use frequency of the verbs, rather than the VPCs, since many of our features are based on the verb of the expression, and moreover, VPC frequency is approximate.) The verification data is used in exploration of the feature space and selection of final features to use in testing; the test set is held out for final testing of the classifiers.

Each VPC in each dataset is annotated by the two human judges according to which of the four senses of *up* identified in Section 3.2 is contributed to the VPC. As noted in Section 1, VPCs may be ambiguous with respect to their particle sense. Since our task here is type classification, the judges identify the particle sense of a VPC in its predominant usage, in their assessment. The observed inter-annotator agreement is 0.80 for each dataset. The unweighted observed kappa scores are 0.73, 0.64 and 0.55, for the training, verification and test sets respectively.

4.2 Calculation of the Features

We extract our features from the 100M word British National Corpus (BNC, Burnard, 2000). VPCs are identified using a simple heuristic based on part-of-speech tags, similar to one technique used by Baldwin (2005). A use of a verb is considered a VPC if it occurs with a particle (tagged AVP) within a six word window to the right. Over a random sample of 113 VPCs thus extracted, we found 88% to be true VPCs, somewhat below the performance of Baldwin’s (2005) best extraction method, indicating potential room for improvement.

The slot and particle features are calculated using a modified version of the ExtractVerb software provided by Joanis and Stevenson (2003), which runs over the BNC pre-processed using Abney’s (1991) Cass chunker.

To compute the word co-occurrence features (WCFs), we first determine the relative frequency of all words which occur within a five word window left and right of any of the target expressions in the training data. From this list we eliminate the most frequent 1% of words as a stoplist and then use the next n most frequent words as “feature words”. For each “feature word”, we then calculate its relative frequency of occurrence within the same five word window of the target expres-

Sense Class	#VPCs in Sense Class		
	Train	Verification	Test
<i>Vert-up</i>	24	33	27
<i>Goal-up</i>	1	1	3
<i>Cmpl-up</i>	20	23	22
<i>Refl-up</i>	15	3	8

Table 1: Frequency of items in each sense class.

Sense Class	#VPCs in Sense Class		
	Train	Verification	Test
<i>Vert-up</i>	24	33	27
<i>Goal-up + Cmpl-up</i>	21	24	25
<i>Refl-up</i>	15	3	8

Table 2: Frequency of items in each class for the 3-way task.

sions in all datasets. We use $n = 200$ and $n = 500$ to create feature sets WCF_{200} and WCF_{500} respectively.

4.3 Experimental Classes

Table 1 shows the distribution of senses in each dataset. Each of the training and verification sets has only one VPC corresponding to *Goal-up*. Recall that *Goal-up* shares a schema with *Cmpl-up*, and is therefore very close to it in meaning, as indicated spatially in Figure 4. We therefore merge *Goal-up* and *Cmpl-up* into a single sense, to provide more balanced classes.

Since we want to see how our features perform on differing granularities of sense classes, we run each experiment as both a 3-way and 2-way classification task. In the 3-way task, the sense classes correspond to the meanings *Vert-up*, *Goal-up* merged with *Cmpl-up* (as noted above), and *Refl-up*, as shown in Table 2. In the 2-way task, we further merge the classes corresponding to *Goal-*

Sense Class	#VPCs in Sense Class		
	Train	Verification	Test
<i>Vert-up</i>	24	33	27
<i>Goal-up + Cmpl-up + Refl-up</i>	36	27	33

Table 3: Frequency of items in each class for the 2-way task.

up/Cmpl-up with that of *Refl-up*, as shown in Table 3. We choose to merge these classes because (as illustrated in Figure 4) *Refl-up* is a sub-sense of *Goal-up*, and moreover, all three of these senses contrast with *Vert-up*, in which increase along a vertical axis is the salient property. It is worth emphasizing that the 2-way task is not simply a classification between literal and non-literal *up*—*Vert-up* includes extensions of *up* in which the increase along a vertical axis is metaphorical.

4.4 Evaluation Metrics and Classifier Software

The variation in the frequency of the sense classes of *up* across the datasets makes the true distribution of the classes difficult to estimate. Furthermore, there is no obvious informed baseline for this task. Therefore, we make the assumption that the true distribution of the classes is uniform, and use the chance accuracy $1/C$ as the baseline (where C is the number of classes—in our experiments, either 2 or 3). Accordingly, our measure of classification accuracy should weight each class evenly. Therefore, we report the average per class accuracy, which gives equal weight to each class.

For classification we use LIBSVM (Chang and Lin, 2001), an implementation of a support-vector machine. We set the input parameters, cost and gamma, using 10-fold cross-validation on the training data. In addition, we assign a weight of $\frac{|Largest\ Class|}{|Class\ c|}$ to each class c to eliminate the effects of the variation in class size on the classifier.

Note that our choice of accuracy measure and weighting of classes in the classifier is necessary given our assumption of a uniform random baseline. Since the accuracy values we report incorporate this weighting, these results cannot be compared to a baseline of always choosing the most frequent class.

5 Experimental Results

We present experimental results for both Ver(ification) and unseen Test data, on each set of features, individually and in combination. All experiments are run on both the 2-way and 3-way sense classification, which have a chance baseline of 50% and 33%, respectively.

Features	3-way Task		2-way Task	
	Ver	Test	Ver	Test
Slots	41	51	53	67
Particles	37	33	65	47
Slots + Particles	54	54	59	63

Table 4: Accuracy (%) using linguistic features.

5.1 Experiments Using the Linguistic Features

The results for experiments using the features that capture semantic and syntactic properties of verbs and VPCs are summarized in Table 4, and discussed in turn below.

5.1.1 Slot Features

Experiments using the slot features alone test whether features that tap into semantic information about a verb are sufficient to determine the appropriate sense class of a particle when that verb combines with it in a VPC. Although accuracy on the test data is well above the baseline in both the 2-way and 3-way tasks, for verification data the increase over the baseline is minimal. The class corresponding to sense *Refl-up* in the 3-way task is relatively small, which means that a small variation in classification on these verbs may lead to a large variation in accuracy. However, we find that the difference in accuracy across the datasets is not due to performance on VPCs in this sense class. Although these features show promise for our task, the variation across the datasets indicates the limitations of our small sample sizes.

5.1.2 Particle Features

We also examine the performance of the particle features on their own, since to the best of our knowledge, no such features have been used before in investigating VPCs. The results are disappointing, with only the verification data on the 2-way task showing substantially higher accuracy than the baseline. An analysis of errors reveals no consistent explanation, suggesting again that the variation may be due to small sample sizes.

5.1.3 Slot + Particle Features

We hypothesize that the combination of the slot features with the particle features will give an increase in performance over either set of linguistic features used individually, given that they tap into differing properties of verbs and VPCs. We find that the combination does indeed give more

Features	3-way Task		2-way Task	
	Ver	Test	Ver	Test
WCF ₂₀₀	45	42	59	51
WCF ₅₀₀	38	34	55	48

Table 5: Accuracy (%) using WCFs.

consistent performance across verification and test data than either feature set used individually. We analyze the errors made using slot and particle features separately, and find that they tend to classify different sets of verbs incorrectly. Therefore, we conclude that these feature sets are at least somewhat complementary. By combining these complementary feature sets, the classifier is better able to generalise across different datasets.

5.2 Experiments Using WCFs

Our goal was to compare the more knowledge-rich slot and particle features to an alternative feature set, the WCFs, which does not rely on linguistic analysis of the semantics and syntax of verbs and VPCs. Recall that we experiment with both 200 feature words, WCF₂₀₀, and 500 feature words, WCF₅₀₀, as shown in Table 5. Most of the experiments using WCFs perform worse than the corresponding experiment using all the linguistic features. It appears that the linguistically motivated features are better suited to our task than simple word context features.

5.3 Linguistic Features and WCFs Combined

Although the WCFs on their own perform worse than the linguistic features, we find that the linguistic features and WCFs are at least somewhat complementary since they tend to classify different verbs incorrectly. We hypothesize that, as with the slot and particle features, the different types of information provided by the linguistic features and WCFs may improve performance in combination. We therefore combine the linguistic features with each of the WCF₂₀₀ and WCF₅₀₀ features; see Table 6. However, contrary to our hypothesis, for the most part, the experiments using the full combination of features give accuracies the same or below that of the corresponding experiment using just the linguistic features. We surmise that these very different types of features—the linguistic features and WCFs—must be providing conflicting rather than complementary information to the classifier, so that no improvement is attained.

Features	3-way Task		2-way Task	
	Ver	Test	Ver	Test
Combined ₂₀₀	53	45	63	53
Combined ₅₀₀	54	46	65	49

Table 6: Accuracy (%) combining linguistic features with WCFs.

5.4 Discussion of Results

The best performance across the datasets is attained using all the linguistic features. The linguistically uninformed WCFs perform worse on their own, and do not consistently help (and in some cases hurt) the performance of the linguistic features when combined with them. We conclude then that linguistically based features are motivated for this task. Note that the features are still quite simple, and straightforward to extract from a corpus—i.e., linguistically informed does not mean expensive (although the slot features do require access to chunked text).

Interestingly, in determining the semantic nearest neighbor of German particle verbs, Schulte im Walde (2005) found that WCFs that are *restricted* to the arguments of the verb outperform simple window-based co-occurrence features. Although her task is quite different from ours, similarly restricting our WCFs may enable them to encode more linguistically-relevant information.

The accuracies we achieve with the linguistic features correspond to a 30–31% reduction in error rate over the chance baseline for the 3-way task, and an 18–26% reduction in error rate for the 2-way task. Although we expected that the 2-way task may be easier, since it requires less fine-grained distinctions, it is clear that combining senses that have some motivation for being treated separately comes at a price.

The reductions in error rate that we achieve with our best features are quite respectable for a first attempt at addressing this problem, but more work clearly remains. There is a relatively high variability in performance across the verification and test sets, indicating that we need a larger number of experimental expressions to be able to draw firmer conclusions. Even if our current results extend to larger datasets, we intend to explore other feature approaches, such as word co-occurrence features for specific syntactic slots as suggested above, in order to improve the performance.

6 Related Work

The semantic compositionality of VPC types has recently received increasing attention. McCarthy et al. (2003) use several measures to automatically rate the overall compositionality of a VPC. Bannard (2005), extending work by Bannard et al. (2003), instead considers the extent to which the verb and particle each contribute semantically to the VPC. In contrast, our work assumes that the particle of every VPC contributes compositionally to its meaning. We draw on cognitive linguistic analysis that posits a rich set of literal and metaphorical meaning possibilities of a particle, which has been previously overlooked in computational work on VPCs.

In this first investigation of particle meaning in VPCs, we choose to focus on type-based classification, partly due to the significant extra expense of manually annotating sufficient numbers of tokens in text. As noted earlier, though, VPCs can take on different meanings, indicating a shortcoming of type-based work. Patrick and Fletcher (2005) classify VPC tokens, considering each as compositional, non-compositional or not a VPC. Again, however, it is important to recognize *which* of the possible meaning components is being contributed. In this vein, Uchiyama et al. (2005) tackle token classification of Japanese compound verbs (similar to VPCs) as aspectual, spatial, or adverbial. In the future, we aim to extend the scope of our work, to determine the meaning of a particle in a VPC token, along the lines of our sense classes here. This will almost certainly require semantic classification of the verb token (Lapata and Brew, 2004), similar to our approach here of using the semantic class of a verb type as indicative of the meaning of a particle type.

Particle semantics has clear relations to preposition semantics. Some research has focused on the sense disambiguation of specific prepositions (e.g., Alam, 2004), while other work has classified preposition tokens according to their semantic role (O'Hara and Wiebe, 2003). Moreover, two large lexical resources of preposition senses are currently under construction, The Preposition Project (Litkowski, 2005) and PrepNet (Saint-Dizier, 2005). These resources were not suitable as the basis for our sense classes because they do not address the range of metaphorical extensions that a preposition/particle can take on, but future work may enable larger scale studies of the type

needed to adequately address VPC semantics.

7 Conclusions

While progress has recently been made in techniques for assessing the compositionality of VPCs, work thus far has left unaddressed the problem of determining the particular meaning of the components. We focus here on the semantic contribution of the particle—a part-of-speech whose semantic complexity and range of metaphorical meaning extensions has been largely overlooked in prior computational work. Drawing on work within cognitive linguistics, we annotate a set of 180 VPCs according to the sense class of the particle *up*, our experimental focus in this initial investigation. We develop features that capture linguistic properties of VPCs that are relevant to the semantics of particles, and show that they outperform linguistically uninformed word co-occurrence features, achieving around 20–30% reduction in error rate over a chance baseline. Areas of on-going work include development of a broader range of features, consideration of methods for token-based semantic determination, and creation of larger experimental datasets.

References

- S. Abney. 1991. Parsing by chunks. In R. Berwick, S. Abney, and C. Tenny, editors, *Principle-Based Parsing: Computation and Psycholinguistics*, p. 257–278. Kluwer Academic Publishers.
- Y. S. Alam. 2004. Decision trees for sense disambiguation of prepositions: Case of over. In *HLT-NAACL 2004: Workshop on Computational Lexical Semantics*, p. 52–59.
- T. Baldwin. 2005. The deep lexical acquisition of English verb-particle constructions. *Computer Speech and Language, Special Issue on Multiword Expressions*, 19(4):398–414.
- C. Bannard. 2005. Learning about the meaning of verb-particle constructions from corpora. *Computer Speech and Language, Special Issue on Multiword Expressions*, 19(4):467–478.
- C. Bannard, T. Baldwin, and A. Lascarides. 2003. A statistical approach to the semantics of verb-particles. In *Proceedings of the ACL-2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, p. 65–72.
- D. Bolinger. 1971. *The Phrasal Verb in English*. Harvard University Press.

- L. Burnard. 2000. *The British National Corpus Users Reference Guide*. Oxford University Computing Services.
- C.-C. Chang and C.-J. Lin. 2001. *LIBSVM: a library for support vector machines*. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- A. Fazly, R. North, and S. Stevenson. 2005. Automatically distinguishing literal and figurative usages of highly polysemous verbs. In *Proceedings of the ACL-2005 Workshop on Deep Lexical Acquisition*.
- B. Fraser. 1976. *The Verb-Particle Combination in English*. Academic Press.
- B. Hampe. 2000. Facing up to the meaning of 'face up to': A cognitive semantico-pragmatic analysis of an English verb-particle construction. In A. Foolen and F. van der Leek, editors, *Constructions in Cognitive Linguistics. Selected Papers from the fifth International Cognitive Linguistics Conference*, p. 81–101. John Benjamins Publishing Company.
- R. Jackendoff. 2002. English particle constructions, the lexicon, and the autonomy of syntax. In N. Dehe, R. Jackendoff, A. McIntyre, and S. Urban, editors, *Verb-Particle Explorations*. Mouton de Gruyter.
- E. Joanis and S. Stevenson. 2003. A general feature space for automatic verb classification. In *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics (EACL-2003)*, p. 163–170.
- R. W. Langacker. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*, volume 1. Stanford University Press, Stanford.
- M. Lapata and C. Brew. 2004. Verb class disambiguation using informative priors. *Computational Linguistics*, 30(1):45–73.
- D. Lin. 1999. Automatic identification of non-compositional phrases. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, p. 317–324.
- S. Lindner. 1981. *A lexico-semantic analysis of English verb particle constructions with out and up*. Ph.D. thesis, University of California, San Diego.
- K. C. Litkowski. 2005. The Preposition Project. In *Proceedings of the Second ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*.
- D. McCarthy, B. Keller, and J. Carroll. 2003. Detecting a continuum of compositionality in phrasal verbs. In *Proceedings of the ACL-SIGLEX Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*.
- D. McCarthy, R. Koeling, J. Weeds, and J. Carroll. 2004. Finding predominant word senses in untagged text. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics*, p. 280–287.
- A. McIntyre. 2001. *The particle verb list*. <http://www.uni-leipzig.de/~angling/mcintyre/pv.list.pdf>.
- P. S. Morgan. 1997. Figuring out *figure out*: Metaphor and the semantics of the English verb-particle construction. *Cognitive Linguistics*, 8(4):327–357.
- T. O'Hara and J. Wiebe. 2003. Preposition semantic classification via Penn Treebank and FrameNet. In *Proceedings of CoNLL-2003*, p. 79–86.
- J. Patrick and J. Fletcher. 2005. Classifying verb-particle constructions by verb arguments. In *Proceedings of the Second ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their use in Computational Linguistics Formalisms and Applications*, p. 200–209.
- P. Saint-Dizier. 2005. PrepNet: a framework for describing prepositions: Preliminary investigation results. In *Proceedings of the Sixth International Workshop on Computational Semantics (IWCS'05)*, p. 145–157.
- S. Schulte im Walde. 2005. Exploring features to identify semantic nearest neighbours: A case study on German particle verbs. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing*.
- K. Uchiyama, T. Baldwin, and S. Ishizaki. 2005. Disambiguating Japanese compound verbs. *Computer Speech and Language, Special Issue on Multiword Expressions*, 19(4):497–512.
- A. Villavicencio. 2005. The availability of verb-particle constructions in lexical resources: How much is enough? *Computer Speech and Language, Special Issue on Multiword Expressions*, 19(4):415–432.
- S. Wurmbrand. 2000. The structure(s) of particle verbs. Master's thesis, McGill University.