Detecting Compositionality of English Verb-Particle Constructions using Semantic Similarity

Su Nam Kim and Timothy Baldwin
Department of Computer Science and Software Engineering
University of Melbourne, VIC 3010, Australia
{snkim,tim}@csse.unimelb.edu.au

Abstract

We present a novel method for detecting the compositionality of English verb-particle constructions (VPCs), based on the assumption that compositionality can be modelled with semantic similarity between VPCs and their base verb in isolation. We also evaluate the contribution of components in compositional VPCs using semantic overlap.

1 Introduction

There has been a recent surge of interest in the compositionality of multiword expressions (hereafter, MWEs), across a range of languages and constructions (McCarthy et al., 2003; Bannard et al., 2003; Baldwin et al., 2003; Uchiyama et al., 2005; Venkatapathy and Joshi, 2005). For the purposes of this paper, compositionality is defined to be a measure of the degree of correspondence between the semantics of the overall MWE and that of its parts. For example, *put up* (as in *put up the painting*) is compositional as it both involves the action of *putting* and the orientation of *up*. *Make up* (as in *kiss and make up*), on the other hand, doesn't involve *making* in any conventional sense of the word, and the semantic contribution of *up* is less than clear, indicating that it is non-compositional.

In this study, we investigate an automatic method for modelling the compositionality of English verb-particle constructions (VPCs). With VPCs, compositionality relates specifically to the semantic contribution of the head verb and particle—which we will refer to as the *components* of the VPC—as contrasted with the overall VPC. Modelling the compositionality of VPCs has applications in lexicography, in determining whether a VPC is non-compositional and should therefore be lexicalized, with a specialised definition, in the dictionary. Another application is parsing, as compositionality has been shown to impact on word order (Dehé, 2002). Villavicencio (2005) noted that given a compositional VPC, it is often possible to predict novel verb-particle combinations from semantically-homogeneous verbs and particles with similar semantics. This observation has been used to classify particle semantics in VPCs (Cook and Stevenson, 2006), and also provides the motivation for our work. That is, when the verb and particle in a given VPC are similar to those in a VPC which is known to be compositional, it is highly likely that the first VPC is compositional. Hence, we can model the compositionality of unseen VPCs via semantic similarity.

2 Motivation

Villavicencio (2005) claimed that verb–particle combinations share the same or similar semantics when the verbs and particles are similar. To detect the compositionality of VPCs, we extend this observation. When two sets of verb–particle combinations (i.e. VPCs) are semantically the same or similar, then their compositionality is also the same. Hence, by looking at the semantic similarity between unseen VPCs and seen VPCs whose compositionality is already resolved, we can predict the compositionality of unseen VPCs. For example, when *call up* is compositional, we can predict that *ring up* is also compositional due to the similarity between *call* and *ring*.

Figure 1 describes the process of predicting the compositionality of the VPC *sum up*. Here, VPCs are grouped by semantics and each group is tagged with its relative compositionality. For example, *put on* and *take on* in `verb class` *n* are semantically similar. Note that the same VPC can vary in its relative compositionality across different interpretations (c.f. *make up the answer* vs. *kiss and make up*). For the purposes of this research, however, we make the (over-)simplifying assumption

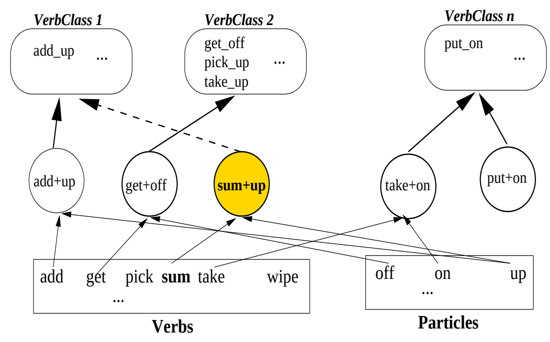

Figure 1: Verb Classes for verb-particle constructions

VPC	compositionality	contribution
lie down	YES	both
close off	YES	verb
get down	YES	particle
make out	NO	none

Table 1: Verb-particle constructions with compositionality and the semantic contribution of its components

that a given VPC has fixed compositionality irrespective of word senses.

In Figure 1, to determine the compositionality of the verb–particle combination *sum up*, we classify it into a semantic class based on lexical similarity. Since *sum* is a synonym of *add*, we predict that *sum up* is in the same semantic class as *add up*, which is compositional. Therefore, we predict that *sum up* is also compositional.

In this process, we need to semantically classify a given combination of verb and particle. To do this, we consider whether each of the verb and particle contribute to the overall semantics of the VPC, leading to four possibilities: verb contribution only, particle contribution only, both verb and particle contribution, or no semantic contribution (similarly to Bannard et al. (2003)).

Examples in Table 1 show the relative semantic contribution of the components in determining the semantics of the overall VPC, and the compositionality of the VPC. As the semantics of the first three VPC examples are derived from the semantics of some or all of its parts, the VPC is compositional. On the other hand, the last example, *make out* is non-compositional since its semantics is not derived from any of its components.

The examples in Table 1 are retrieved from the two available lexical resources: Bannard (2006) and McCarthy et al. (2003). The data is described in detail in Section 4.

Finally, we claim that the semantic similarity of VPCs can be used as a predictor of VPC compositionality. We look at the semantic similarity by checking the semantic contribution of the VPC components. To semantically classify a given combination of verb and particle, we use WORDNET 2.1 hypernyms as the semantic classes of the verb. Note that in our experiment, we use two different forms of hypernyms: direct hypernyms (the first hypernym) and root hypernyms (the *N*th hypernym). For particle semantics, we adopt the classification of Bannard et al. (2003) described in Section 4. To detect the semantic contribution, we build a classifier that takes the semantic class of the VPC, verb and particle as features.

3 Related Work

Lin (1999) used distributional similarity to predict the compositionality of a range of multiword expressions (MWEs), and Schone and Jurafsky (2001) used distributional similarity between the multiword expression candidates and their component words. Both of these works focused primarily on noun–noun MWEs. Baldwin et al. (2003) used a statistical model (in the form of latent semantic analysis) to test the compositionality of VPCs and noun–noun compounds. Detecting the compositionality of VPCs have also been attempted by Bannard et al. (2003) based on the contribution of the verb and particle in a given VPC to the semantics of the VPC. In this work, Bannard et al. (2003) experimented with four tasks: detecting compositionality, one component (either verb or particle) contributes the semantics, verb contributes the semantics, and particle contributes the semantics. McCarthy et al. (2003) also used distributional semantics to predict the compositionality of VPCs, based on the relative similarity between the verb and VPC. McCarthy et al. (2003) modelled similarity via a number of variants of the overlap of the top *N* neighbours (based on distributional similarity). They also carried out comparative evaluation with a thesaurus, man-made resources and collocation measures. Venkatapathy and Joshi (2005) claimed that if MWEs are decomposable, then they are more likely to be compositional. Piao et al. (2006) tested the compositionality of MWEs using a semantic field taxonomy based on the Lancaster English semantic lexicon.

In this paper, we are exclusively interested in the compositionality of English verb-particle constructions (VPCs). There is a separate but related body of computational linguistic research has been done on the extraction and identification of VPCs. Baldwin and Villavicencio (2002) and Baldwin (2005) extracted VPCs from untagged text using syntactic features. Kim and Baldwin (2006) identified VPCs using an assortment of syntactic and semantic features. O'Hara and Wiebe (2003) proposed a method for preposition sense disambiguation in verb–preposition combinations.

4 Data Collection

We used data from two sources in our experiments: data from Bannard (2006), which individually classifies the compositionality of the verb and particle in VPCs; and data from McCarthy et al. (2003)[1], where the overall VPC is manually scored for compositionality. We describe these two resources in detail below.

4.1 Bannard (2006) Dataset

Bannard (2006) manually tagged 160 VPCs for verb and particle compositionality. The tagging involved 29 human annotators who made a binary judgement for whether first the verb then the particle contributes semantics to the overall VPC. If they could not decide on the contribution of the verb or particle, a tag of "don't know" was assigned. The VPCs in the data set are considered **compositional** if more than half of human annotators answered "yes" for the verb, particle or both. We also generated verb-specific and particle-specific compositionality judgements, giving rise to a total of three datasets: **verb-compositional**, **particle-compositional**, and **VP-compositional**. Table 3 describes the number of compositional and non-compositional VPCs and verbs in isolation in the data set.

Note that in Table 3, the totals are less than the number in the original dataset as a result of filtering out all VPCs where the VPC or base verb was not found in WordNet 2.1.

[1]Available from mwe.stanford.edu/resources/

	VPC	Verb*
Verb-compositional	52	59
Particle-compositional	12	11
VP-compositional	50	62
Non-Compositional	22	25
TOTAL	136	158

Table 2: Breakdown of data from Bannard (2006)

	VPC	Verb*
Compositional	35	58
Non-Compositional	32	39
TOTAL	67	97

Table 3: Breakdown of data from McCarthy et al. (2003)

4.2 McCarthy et al. (2003) Dataset

McCarthy et al. (2003) manually scored 117 VPCs in terms of overall compositionality using a scale of 0 to 10, where 10 indicates the VPC is fully compositional and 0 indicates it is fully non-compositional. For our purposes, we binarised this data such VPCs with a score of 5 or higher are considered to be compositional and all other VPCs are considered to be non-compositional.

Due to the lack of instances in WordNet 2.1, we finally extracted 67 VPCs and 97 verbs in isolation from McCarthy et al. (2003).

4.3 Semantics of Particles

In order to establish a means to quantify the lexical similarity of particles in VPCs, we classified particles in VPCs according to the following 4-way categorisation (based on Bannard (2003)): TEMPORAL, SPATIAL-DIRECTION, SPATIAL-POSITION and NONE (the last of which is used to tag non-compositional particles). Two human annotators tagged the particles in the VPCs across the two datasets. The initial inter-annotator agreement for this task was 51.71%, and all disagreements were subsequently resolved between the annotators. The final distribution of particles was 86 SPATIAL-DIRECTION, 49 SPATIAL-POSITION and 121 NONE (for a total of 256 VPCs).[2]

[2]Note that there were no instances of TEMPORAL particles in this dataset, as TEMPORAL particles tend to be fully productive and the VPC thus doesn't occur in WordNet 2.1,

We observed that the semantics of particles do not always match the contribution of the particle in the VPCs. Among the 74 VPCs where the particle was found to be compositional (either particle-compositional or VP-compositional), the particle was tagged as NONE in 22 cases. This discrepancy is due to the fact that we have binarised the compositionality data (by taking the majority decision in the case of the Bannard et al. (2003) data, and applying a threshold of 0.5 for the McCarthy et al. (2003) data), generating a mix of fully and partially compositional VPCs. The partially compositional VPCs cause the problems. For example, in the Bannard et al. (2003) data, 14 annotators marked *knock out* as particle-compositional and 12 annotators marked it as particle-non-compositional. It is thus considered to be particle-compositional for our purposes, but our annotators agreed that the most appropriate semantic class for the particle was NONE.

5 Classifier Construction

We experimented with two approaches to compositionality classification. The first approach (**C1**) models compositionality in the two datasets via the verb and VPC semantics according to WORDNET 2.1, the hand-annotated particle semantics verb–particle co-occurrence data. In the approach, we model compositionality via a verb–particle matrix representation of VPC attestation, based on the research of Uchiyama et al. (2005) over Japanese compound verbs (**C2**). In each case, we experimented with TIMBL (Daelemans et al., 2004) and Zhang Le's MAXENT toolkit.[3]

Uchiyama et al. (2005) successfully performed a Japanese compound verb semantic classification task based on a 2-dimensional matrix representation of which main and auxiliary verb (e.g. *yuderu* "boil" and *ageru* "up") combine to form a compound verb (*yude-ageru* "boil up", in this case). The rows in the matrix represent which auxiliary verb a given main verb combines with to form a compound verb. The feature representation for a given V1-V2 combination is made up of the concatenation of the corresponding row and column in the matrix. Uchiyama et al. (2005) experimented with two basic feature–value representations: (1) simple co-occurrence, i.e. is the given V1-V2 combination attested in corpus data, and (2) semantic categorisation, i.e. is the V1-V2 combination attested in training data, and if so, what is the semantic class associated with it.

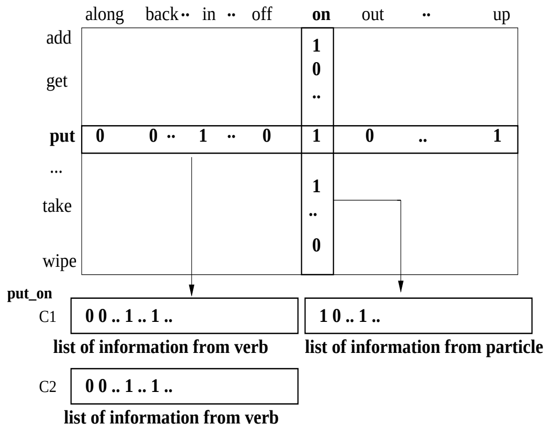

Figure 2: Collocational feature representation

Since the combinatoric properties of English VPCs are similar to those of Japanese compound verbs, this methodology is directly applicable to the compositionality task as targeted in this research. We adapt the two feature–value representations from Uchiyama et al. (2005) in combing up with the following four possibilities: (1) **C–C** where both the verb and particle are represented by way of simple (binary) co-occurrence; (2) **C–S** where the verb is represented via co-occurrence with different particles, but the particle is represented by its semantics (e.g. SPATIAL-POSITION, or 0 if it doesn't produce a VPC in combination with the given verb); (3) **S–C** where the verb is represented via semantics (i.e. WORDNET 2.1 synsets) and the particle via simple co-occurrence; and (4) *S–S* where both the verb and particle are represented via semantics. The input features of two classifiers, (C1) and (C2) are described in Table 4.

The VPC class labels in each case depend on the dataset. For the McCarthy et al. (2003) dataset, the classes are YES and NO, indicating compositionality and non-compositionality, respectively. For the Bannard (2006) dataset, the classes are VERB, PARTICLE, BOTH and NONE.

In Figure 2, (C2) is constructed by concatenating the corresponding row and column of the matrix. The feature–value for each can take the form

e.g. *eat up*. If we removed the restriction on the VPC being lexicalized in WORDNET 2.1, we would expect TEMPORAL VPCs to make up a significant proportion of the data.

[3] http://homepages.inf.ed.ac.uk/s0450736/maxent_toolkit.html

Test	Train	Experiment ID	Classifier Type	Features
VPC	VPC	E1	C1	V_{SEM}, V_{SEM}+P, V_{SEM}+P+P_{SEM}
		E2	C2	C–C, C–S, S–C, S–S
V	V	E3	C1	V_{SEM}, V_{SEM}+P, V_{SEM}+P+P_{SEM}
V	VPC	E4	C1	V_{SEM}, V_{SEM}+P, V_{SEM}+P+P_{SEM}

Table 4: Experimental setup (V_{SEM}= verb semantics, P = particle wordform and P_{SEM}= particle semantics; C = binary co-occurrence and S = semantic class)

Experiment	Features	Semantics	Compositionality	$\mathscr{P}$	$\mathscr{R}$	$\mathscr{F}$
E1	baseline	–	Y	.731	–	–
		–	N	.269	–	–
	V	root	Y	.743	.728	.736
			N	.298	**.315**	.307
E2	baseline	1st, root	Y	.761	–	–
			N	.239	–	–
	S–C	1st	Y	**.791**	.895	**.843**
			N	.429	.250	**.339**
E3	baseline	–	Y	.749	–	–
		–	N	.251	–	–
	V	1st	Y	.771	.864	.818
			N	.366	.234	.300
E4	baseline	–	Y	.598	–	–
		–	N	.402	–	–
	V	1st	Y	.632	**.948**	.759
			N	**.700**	.180	.286

Table 5: Results of experiments with TiMBL (the best result in each column for each of compositional and non-compositional is indicated in **boldface**)

of either simple co-occurrence or semantics, giving rise to the four combinations described in Section 6.1.

6 Evaluation

6.1 Evaluation Set

The underlying hypothesis in this research is that compositionality can be modelled by the lexical semantics of the verb and particle relative to that of the overall VPC. The verb and VPC lexical semantics are represented by WordNet 2.1 synsets, along with the direct hypernyms and root hypernyms (i.e. unique beginner ancestor) of each. We then represent the semantics of the particles using the 4-way semantic classification described in Section 4.3. Additionally, we have two different classifier structures, (C1) and (C2), based on the simple VPC vs. the collocational properties of the verb and particle in the VPC. As a result, we have 25 different sets of data for the test.

Of the 4 experiments outlined in Table 4 (E1–E4), we used 10-fold cross-validation for E1–E3 and a fixed partitioning of the data for E4, due to the asymmetry in training and test data (verb vs. VPC semantics, respectively).

6.2 Detecting Compositionality

We used two different learners, namely TiMBL and MaxEnt, to build our various classifiers. For each experiment, we present a baseline majority-class classifier, along with only the best performing classifier due to space limitations. Evaluation is in terms of precision ($\mathscr{P}$), recall ($\mathscr{R}$) and F-score ($\mathscr{F}$) initially, and then classification accuracy ($\mathscr{A}$) and Pearson correlation (ρ), following McCarthy et al. (2003).

We present the results using TiMBL in Table 5 and those using MaxEnt in Table 6. Our classification method performed well when using the semantics of both the VPC and the verb. On the con-

Experiment	Features	Semantics	Compositionality	$\mathscr{P}$	$\mathscr{R}$	$\mathscr{F}$
E1	baseline	–	Y	.731	–	–
		–	N	.269	–	–
	VPS	root	Y	.751	.905	.828
			N	.435	.185	.310
E2	baseline	C–C	Y	.761	–	–
			N	.239	–	–
	C–S	1st	Y	.759	.993	**.876**
			N	.000	.000	.000
E3	baseline	–	Y	.749	–	–
		–	N	.251	–	–
	VPS	1st	Y	**.773**	.895	.836
			N	.429	**.234**	.332
E4	baseline	–	Y	.598	–	–
		–	N	.402	–	–
	VP,VPS	1st	Y	.604	**1.00**	.802
			N	**1.00**	.026	**.513**

Table 6: Results of experiments with MAXENT (the best result in each column for each of compositional and non-compositional is indicated in **boldface**)

trary, while the results are not included in this paper, in general, particle semantics decreased classifier performance due to a lack of data. Also, of the two WORDNET 2.1 sense generalisation methods (1st vs. root hypernyms), we found that 1st hypernyms classified verbs better since they preserve the specific meaning of verbs better.

Comparing the two classifier strategies (C1) and (C2), our results indicate that (C2) produced the best performance among all experiments from both TIMBL and MAXENT, mirroring the results of Uchiyama et al. (2005) for Japanese compound verbs. The highest F-score for TIMBL and MAXENT at detecting compositional VPCs was 0.843 and 0.876, respectively. The results for non-compositional VPC detection were considerably lower, with (C1) producing the highest F-score at 0.513 in experiment 4 using MAXENT. This can be explained by a relative sparsity of non-compositional VPCs in the dataset. Note, the results of all our classifiers for both TIMBL and MAXENT exceeded the baseline in all cases.

6.3 Detecting Semantic Contribution

In this secondary experiment, we test the semantic contribution of the verb and particle in compositional VPCs. The data for this experiment is described in Table 7.

In order to evaluate our approach, we built three classifiers using TIMBL and 10-fold cross-validation: (1) two classifiers based on (C1), with either VPC or verb semantics; and (2) a classifier based on (C2). As shown in Table 8, the highest achieved accuracy of 81.95% was with (C2) using the verb and particle semantics (S–S) and the 1st hypernym of the verb and VPC. Our classifier outperformed the baseline in terms of Precision except for the particle-compositionality task, due to data sparseness.

Feature	Comp	Non-C	total
VPC	112	22	134
	Verb (52)		
	Particle (10)		
	Both (50)		
V	133	25	158
	Verb (59)		
	Particle (62)		
	Both (12)		

Table 7: Number of instances for semantic contribution (Comp = compositionality, Non-C = non-compositionality)

Feature	Comp	Cont	$\mathscr{P}$	$\mathscr{R}$	$\mathscr{F}$
baseline	yes	VP	.373	–	–
		V	.388	–	–
		P	.075	–	–
		All	.836	–	–
	no		.164	–	–
S–S	yes	VP	.489	.460	.475
		V	.474	.745	.610
		P	.000	.000	.000
		All	**.843**	**.964**	**.903**
	no		**.333**	**.091**	**.212**

Table 8: Result of contribution of verb and particle (Comp = compositionality, Cont = contribution)

6.4 Measuring Correlation

Finally, for direct comparison with McCarthy et al. (2003), we evaluate the different classifiers using classification accuracy ($\mathscr{A}$). For the MAXENT classifier, we additionally evaluate the Pearson correlation (ρ) between the compositionality probabilities and the original (non-binarised) human judgements. The results are presented in Table 9. In this case, **boldface** figures indicate above-baseline performance.

The range of correlation statistics reported in McCarthy et al. (2003) was between -0.115 and 0.490 whereas ours were between -0.070 and 0.274. Although the best correlation achieved by our models is lower than that of McCarthy et al. (2003) (0.49 vs. 0.27), it is worth reflecting that McCarthy et al. (2003) used distributional similarity over the entire British National Corpus (i.e. a data and computationally intensive approach), as contrasted with our experiments over small amounts of training data but incorporating lexical semantic resources (i.e. a knowledge intensive approach). Out of our classifiers, (C1) using the direct hypernyms as to represent verb semantics (i.e. E3, with V_{SEM}(1st)) produced the highest correlation.

7 Conclusion

In this study, we proposed a method to automatically detect the compositionality of VPCs based on semantic similarity with their component verb and particle. To evaluate our method, we explored two learning approaches, based on the lexical semantics of only the VPC, verb and particle concerned (C1), and verb and particle co-occurrence (C2). In representing the verb and VPC, we experimented with two methods of semantic generalisation, using direct hypernyms (1st) and the unique beginner ancestor (root). We built classifiers using TIMBL and MAXENT, and evaluated our methods both in terms of precision, recall and F-score for a binary compositionality classification task, and correlation between the marginal probability estimates of the MAXENT model and human judgements (on a continuous scale). The performance of our method exceeded the baseline for each class in terms of F-score. We also evaluated the contribution of the verb and particle in compositional VPCs using the same method to predict the individual verb- and particle-compositionality prediction tasks.

8 Acknowledgement

Many thanks to Dr.Colin Bannard and Dr.Diana McCarthy for data and to the three anonymous reviewers for their valuable input.

References

Timothy Baldwin. 2005. The deep lexical acquisition of English verb-particle constructions. *Computer Speech and Language, Special Issue on Multiword Expressions*, 19(4):398–414.

Timothy Baldwin, Colin Bannard, Takaaki Tanaka, and Dominic Widdows. 2003. An empirical model of multiword expression decomposability. In *Proc. of the ACL-2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, pages 89–96. Sapporo, Japan.

Timothy Baldwin and Aline Villavicencio. 2002. Extracting the unextractable: A case study on verb-particles. In *Proc. of the 6th Conference on Natural Language Learning (CoNLL-2002)*, pages 98–104. Taipei, Taiwan. URL `http://lingo.stanford.edu/pubs/tbaldwin/conll2002.pdf`.

Colin Bannard. 2003. *Statistical Techniques for Automatically Inferring the Semantics of Verb-Particle Constructions*. Master's thesis, University of Edinburgh.

Colin Bannard. 2006. *Acquiring Phrasal Lexicons*

Experiment	Semantics(N)	TiMBL $\mathscr{A}$	MAXENT $\mathscr{A}$	ρ
E1	baseline	.731		
	V_{SEM}(1st)	.527	.640	.181
	V_{SEM}(root)	.617	.670	.148
	V_{SEM}+P(1st)	.597	.665	.196
	V_{SEM}+P(root)	.617	.680	.148
	V_{SEM}+P+P_{SEM}(1st)	.547	.680	.127
	V_{SEM}+P+P_{SEM}(root)	.592	.715	.085
E2	baseline(C)	.736		
	baseline(1,2)	.751		
	C-C	.711	**.740**	.100
	C-S(1st)	.706	**.755**	-.070
	C-S(root)	.706	**.755**	-.070
	S-C(1st)	.740	.670	.029
	S-C(root)	.735	.665	.020
	S-S(1st)	.685	.720	.155
	S-S(root)	.660	.720	.152
E3	baseline	.749		
	V_{SEM}(1st)	.706	.703	**.274**
	V_{SEM}(root)	.706	.704	.183
	V_{SEM}+P(1st)	.702	.719	.274
	V_{SEM}+P(root)	.671	.692	.163
	V_{SEM}+P+P_{SEM}(1st)	.671	.731	.250
	V_{SEM}+P+P_{SEM}(root)	.667	.700	.140
E4	baseline	.598		
	V_{SEM}(1st)	**.639**	**.630**	.237
	V_{SEM}(root)	**.619**	.580	.072
	V_{SEM}+P(1st)	**.629**	.588	.204
	V_{SEM}+P(root)	**.608**	**.608**	.121
	V_{SEM}+P+P_{SEM}(1st)	.577	**.608**	.157
	V_{SEM}+P+P_{SEM}(root)	**.608**	**.608**	.106

Table 9: Accuracy and correlation (C = the hypernym level, V_{SEM}= verb semantics, P = particle wordform, P_{SEM}= particle semantics)

from Corpora. Ph.D. thesis, University of Edinburgh.

Colin Bannard, Timothy Baldwin, and Alex Lascarides. 2003. A statistical approach to the semantics of verb-particles. In *Proc. of the ACL-2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, pages 65–72. Sapporo, Japan.

Paul Cook and Suzanne Stevenson. 2006. Classifying particle semantics in English verb-particle constructions. In *Proc. of the ACL-2006 Workshop on Multiword Expressions: Identifying and Exploiting Underlying Properties*, pages 45–53. Sydney, Australia.

Walter Daelemans, Jakub Zavrel, Ko van der Sloot, and Antal van den Bosch. 2004. *TiMBL: Tilburg Memory Based Learner, version 5.1, Reference Guide*. ILK Technical Report 04-02. URL `http://ilk.uvt.nl/downloads/pub/papers/ilk0402.pdf`.

Nicole Dehé. 2002. *Particle Verbs in English: Syntax, Information, Structure and Intonation*. John Benjamins, Amsterdam, Netherlands/Philadelphia USA.

Su Nam Kim and Timothy Baldwin. 2006. Automatic extraction of verb-particles using linguistic features. In *11th Conference of European Chapter of the Association for Computational*

Linguistics : Third ACL-SIGSEM workshop on Preposition. Italy.

Dekang Lin. 1999. Automatic identification of non-compositional phrases. In *In proceedings of ACL-99*, pages 317–324. Maryland.

Diana McCarthy, Bill Keller, and John Carroll. 2003. Detecting a continuum of compositionality in phrasal verbs. In *Proc. of the ACL-2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*. Sapporo, Japan.

Tom O'Hara and Janyce Wiebe. 2003. Preposition semantic classification via Treebank and FrameNet. In *Proc. of the 7th Conference on Natural Language Learning (CoNLL-2003)*, pages 79–86. Edmonton, Canada.

Scott Piao, Paul Rayson, Olga Mudraya, Andrew Wilson, and Roger Garside. 2006. Measuring MWE compositionality using semantic annotation. In *Proc. of the ACL-2006 Workshop on Multiword Expressions: Identifying and Exploiting Underlying Properties*, pages 2–11. Sydney, Australia.

Patrick Schone and Daniel Jurafsky. 2001. Is knowledge-free induction of multiword unit dictionary headwords a solved problem? In *In proceedings of the 2001 conference on Empirical Methods in Natural Language Processing*, pages 100–108. Hong Kong.

Kiyoko Uchiyama, Timothy Baldwin, and Shun Ishizaki. 2005. Disambiguating Japanese compound verbs. *Computer Speech and Language, Special Issue on Multiword Expressions*, 19(4):497–512.

Sriram Venkatapathy and Aravind K. Joshi. 2005. Measuring the relative compositionality of verb-noun (V-N) collocations by integrating features. In *Proc. of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing(HLT/EMNLP 2005)*, pages 899–906. Vancouver, Canada.

Aline 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.