The Semantics of *Backing Up*
(Or: What to do with prepositions and particles?)

Marjorie McShane, Stephen Beale and Sergei Nirenburg

University of Maryland Baltimore County

{marge,sbeale,sergei}@umbc.edu

Abstract

The paper discusses the lexical description and runtime disambiguation of homographous prepositions and verbal particles in English using the Ontological Semantic (OntoSem) text processing environment. It describes the knowledge resources and processors that permit – in all but the most genuinely ambiguous of cases – an unambiguous text-meaning representation (TMR) to be created from input containing such multifunctional elements. The example of *back up* in many of its senses is treated in particular detail.

Introduction

In English, many prepositions are homographous with (i.e., have the same spelling as) verbal particles. We will call this mixed class of elements P-Ps (prepositions/particles). Consider, for example, the collocation *go after + NP*, which can have two different syntactic analyses:

- verb + particle + direct object, with the idiomatic meaning *pursue*: *The cops went after the criminal.*
- verb + preposition + object of preposition, with the compositional meaning do some activity after some other activity is finished: *The bassoonist went after the cellist.*

While there are clearly two syntactic analyses of *go after* that are associated with different meanings, and while there is often a default reading depending on the subject and object selected, one actually cannot confidently select one or the other interpretation outside of context: after all, *The cops went after the criminal* could mean that the cops provided testimony after the criminal finished doing so, and *The bassoonist went after the cellist* could mean that the former pursued the latter for having stepped on his last reed. Examples like these are worst-case scenarios for machine processing systems because they are genuinely ambiguous outside of the broader discourse context. Luckily, it is more common for sentences to have a single interpretation without the need for extra-clausal inferencing. In this paper we present knowledge-based methods for arriving at a single, unambiguous interpretation of many sentences containing P-Ps that, in knowledge-poor environments, would remain ambiguous or uninterpreted.

The text processing approach to be discussed is Ontological Semantics (OntoSem). In this paper we give a brief overview of OntoSem, describe the lexical and ontological knowledge bases that support the disambiguation of P-Ps, describe the analyzer that uses those resources to produce unambiguous text-meaning representations (TMRs) from input text, and place this work in the bigger picture. For reasons of space we orient the discussion around short, invented examples, not the newspaper sentences – whose average length is more than 25 words – that our system usually processes.

A Snapshot of OntoSem

OntoSem is a text-processing environment that takes as input unrestricted raw text and carries out preprocessing, morphological analysis, syntactic analysis, and semantic analysis, with the results of semantic analysis represented as formal text-meaning representations (TMRs) that can then be used as the basis for many applications (for details, see, e.g., Nirenburg and Raskin 2004, Beale et al. 1995, 2003). Text analysis relies on:

- The OntoSem language-independent ontology, which is written using a metalanguage of description and currently contains around 8,500 concepts, each of which is described by an average of 16 properties.
- An OntoSem lexicon for each language processed, which contains, among other information, syntactic (syn-struc) and semantic (sem-struc) zones as well as calls for procedural semantic routines. The semantic zone most frequently refers to ontological concepts, either directly or with property-based modifications, but can also describe word meaning parametrically, for example, in terms of modality, aspect, time, etc. The current English lexicon contains approximately 25,000 senses, including most closed-class items and many of the most frequent (and often difficult) verbs, as indicated by corpus analysis (for a detailed description of the lexicon go to http://ilit.umbc.edu.).
- An onomasticon, or lexicon of proper names, of approximately 350,000 entries.
- A fact repository, which contains real-world facts represented as indexed "remembered instances" of ontological concepts; e.g., SPEECH-ACT-3366 is the 3366th instantiation of the concept SPEECH-ACT in the world

model constructed during the processing of some given text(s).
- The OntoSem text analyzers, which cover preprocessing, syntactic analysis, semantic analysis, and the creation of TMRs. Instead of using a large, monolithic grammar of a language, which leads to ambiguity and inefficiency, we use a special lexicalized grammar created on the fly for each input sentence (Beale, et. al. 2003). Syntactic rules are generated from the lexicon entries of each of the words in the sentence, and supplemented by a small inventory of generalized rules. We augment this basic grammar with transformations triggered by words or features present in the input sentence.
- The TMR language, which is the metalanguage for representing text meaning.

At the present stage of development, the TMR, together with the static (ontology, lexicons, fact repository) and dynamic (analyzer) knowledge sources that are used in generating and manipulating it, already provide mid-size coverage for a variety of semantic phenomena. In fact, the TMR represents, in a compact way, practically attainable solutions for many issues that have concerned the computational linguistics and NLP community for over forty years. Our TMRs have been used as the substrate for question-answering, MT (Beale et al. 1995), knowledge extraction (Beale et al. 2004), and were used as the basis for reasoning in the question-answering system AQUA, where they supplied knowledge to enable the operation of the JTP (Fikes et al., 2003) reasoning module.

Lexical and Ontological Knowledge

The OntoSem lexicon is based on LFG formalism, extended in significant ways to support, for example, calls to procedural semantics programs for the determination of context-based meaning (McShane et al. 2004a).

For practical reasons, we currently distinguish five classes of P-Ps in our lexicon (examples are presented in an abridged format). [1]

1. Productive senses of prepositions: for example, the second prepositional sense of *after* sets up the syntactic expectation the input will contain a clause followed by a prepositional phrase headed by the head word of the entry, *after*. The *sem-struc* zone shows our primarily extra-ontological method of expressing temporal relations.

after-prep2
synonyms "in_the_wake_of" "in_the_aftermath_of"
def "a temporal relation; absolute time, not time span"
ex "I will go to school after the dance."
syn-struc
  cl (root $var1) (cat cl)
  pp (root $var0) (obj ((root $var2) (cat np)))
sem-struc
  ^$var1 (sem EVENT) (time (> (value ^$var2.time)))
  ^$var2 (sem EVENT)

The *syn-struc* and *sem-struc* zones are linked by variables. The meaning of value of each variable in this example is constrained to be EVENT (ontological concepts are written in small caps).

2. Productive senses of prepositions that are used as adverbs with semantic ellipsis (for a discussion of semantic ellipsis, see McShane et al. 2004b). For example, *in-adv1* covers contexts like *He walked in late,* in which the DESTINATION must be computed by procedural semantics using the context of the surrounding sentences. A call to the procedure that carries this out is encoded in the lexical entry for *in-adv1*. Adverbs that are homographous with prepositions are important because they can lead to analysis ambiguity between the adverbial reading and a particle reading (e.g., *walk* in_{ADVERB} = 'enter' versus *give* $in_{PARTICLE}$ = 'submit').

3. Prepositions that head prepositional phrases which are listed in the lexical description of another entity. These divide into three types:

(a) those that must be listed because their meaning is not compositional: e.g., *account for NP* ("bad weather accounted for the cancellation"), where the meaning of *for* cannot be reconstructed apart from the verb;[2]

(b) those that must be listed because the syntactic structure is not valid without a PP of some sort: e.g., the verb *back* can be used in the meaning *move backwards* with many different PPs (*she backed into the tree, around the bush, out of the garage*), but it cannot be used without any PP at all: **She backs/she is backing* (cf. the discussion below); therefore, we list a sense of *back* that takes a PP, but leave the meaning of the PP to be analyzed compositionally at runtime;

(c) those that do not strictly have to be listed because their meaning is compositional, but often are listed in the OntoSem lexicon in order to reduce potential ambiguity that the analyzer might encounter. For example, most deverbal nouns take two optional PP adjuncts headed by *of* and *by*, whose complements generally link to the THEME and AGENT, respectively, of the given event: *certification of the document by the lawyer*. By recording these optional PPs with their typical meanings in the deverbal entries, like the one for *certification*, we tell the analyzer to prefer this interpretation over other compositional meanings of the

[1] Our categorization is more coarse-grained than the one in Huddleston and Pullum (2002). As with all microtheories in Ontological Semantics, the research, knowledge acquisition and implementation of the microtheory of P-Ps is occurring in iterative cycles of refinement, with the goal of continuously producing useful practical applications.

[2] This category conflates several of Huddleston and Pullam's (2002), which are motivated in a large measure by available syntactic transformations.

PPs that otherwise might receive equal scores upon analysis. In the example of *certification*, we want to avoid the equally plausible, but pragmatically much less likely, interpretation of *by* that would result in the interpretation 'certification of the document that is located next to the lawyer'. The analyzer gives preference to lexically encoded PPs over compositional meanings since this solution tends to be correct and massively reduces potential parses. Additional acquisition time is minimal, since acquisition is supported by templates, including one for deverbal nouns.

4. Verbal particles that can be separated from their verbs: these are of the type *block off the road/block the road off; blow up the bridge/blow the bridge up; boot someone out/boot out someone.* We use the syntactic label *prep-part* (prepositional particle) to refer to such particles. This label indicates that the particle is subject to various ordering options, which are handled by special rules in the OntoSem analyzer.

blow-v1
def "phrasal: blow up"
ex "they blew up the bridge"
syn-struc
  subject (root $var1) (cat np)
  (root $var0) (cat v)
  prep-part (root up) (root $var2) (cat prep)
  directobject (root $var3) (cat np))
sem-struc
  EXPLODE
    THEME (value ^$var3)
    CAUSED-BY (value ^$var1)
 ^$var2 (null-sem +)

5. Verbal particles that are not detachable from their verbs because (a) they take no direct object: *She backed up, the talks broke down* or (b) they take a direct object but it cannot intervene: *He banked on a raise (*He banked a raise on), They bought into his philosophy.* Lexical encoding is similar to the *explode* example except that the particle is called *prep-part-fixed*, which asserts fixed word ordering.

Producing Unambiguous TMRs from Input that Includes P-Ps

The OntoSem syntactic parser and semantic analyzer do a good job of disambiguating P-Ps whose meanings are recorded in the static knowledge sources. The parser is able to keep several analyses available until semantics can disambiguate, and it handles discontinuous entities even when a significant amount of text separates them (this is not possible, for example, using N-grams to process multi-word entities). For example, the analyzer generated a single, correct TMR for the following sentence, despite the distance between the verb (*laid*) and its particle (*off*): *The manager laid the employee who wrote the letter to the newspaper off*. The semantic analyzer uses a combination of lexical specifications and ontological knowledge to disambiguate the parse, as well as disambiguate between senses of a given input with the same parse. (In fact, the need for disambiguation has been a driving force in the development of Ontological Semantics; it is discussed at length in Nirenburg and Raskin 2004, Chaper 8 and Beale et al. 1995.)

We describe the processing of P-Ps in OntoSem using a set of related examples (see next page), all of which contain the verb *back* and the P-P *up*, which in (a)-(e) is a verbal particle listed in a corresponding lexicon entry, and in (f) is a preposition that has compositional semantics. We know that the last example is a productive use of the preposition because any directional preposition + NP can be used with *back* in its intransitive sense: *she backed away from the wolf, toward the door, around the garage…*

The TMRs for the examples omit some details, for reasons of space. The numbers associated with ontological concepts indicate the concept instance during this run of the analyzer; we started a new run for this experiment so the numbers are low. Each concept instance (indicated in boldface) heads its own "sub-entry" in the TMR, which specifies the instance's own properties and values, as well as its relationship to other instances in the TMR. For example, in (1a) the concept instance HUMAN-1 is listed both as the AGENT in the sub-entry for CHANGE-LOCATION-1, and in its own subentry, where it listed as the AGENT-OF: CHANGE-LOCATION-1 (in short, there is cross-referencing within the TMR, and an "entity profile" is created for each object and event that occurs in the text).

If we are to generalize about the challenges of automatically generating TMRs from list of sentences like those on the next page, the main problem is ambiguity, both syntactic and semantic. Below we describe OntoSem's automatic means of disambiguating each example in turn.

Examples 1a (*The man backed up*) and 1b (*The car backed up*). Syntactically, these are unambiguously analyzed as [subject–verb–particle]. The only way they could have been analyzed differently is if *back* had had a sense that permitted the structure [subject–verb] with no complements, in which case the productive meaning of the adverb *up* (as in *The bird flew up*) might have been appended to it. However, since there is no such sense of *back* (one cannot say **He backs/He is backing*), the only available interpretation is [subject–verb–particle].

The lexical senses that cover (1a) and (1b) have identical syntactic structures but differ in the semantic constraints on their components and the case-roles they realize. In one sense, the meaning of the subject is constrained to ANIMAL and its case-role is the AGENT of a MOTION-EVENT whose DIRECTION-OF-MOTION is BACKWARD; in the other sense, the meaning of the subject is constrained to VEHICLE and its case-role is THEME of the

1a. The man backed up.

```
CHANGE-LOCATION-1
   (AGENT: HUMAN-1) (DIRECTION-OF-MOTION: BACKWARD)
   (TIME: < (FIND-ANCHOR-TIME)
HUMAN-1
   (GENDER: MALE) (AGE: > 18)
   (AGENT-OF: CHANGE-LOCATION-1)
```

1b. The car backed up.

```
CHANGE-LOCATION-2
   (THEME: AUTOMOBILE-1)
   (DIRECTION-OF-MOTION: BACKWARD)
   (TIME: < (FIND-ANCHOR-TIME)
AUTOMOBILE-1
   (THEME-OF: CHANGE-LOCATION-2)
```

1c. The woman backed up her husband.

```
SUPPORT-1
   (AGENT: HUMAN-3) (BENEFICIARY: SPOUSE-1)
   (TIME: < FIND-ANCHOR-TIME)
HUMAN-3
  (GENDER: FEMALE) (AGE: > 18) (AGENT-OF: SUPPORT-1)
SPOUSE-1
  (GENDER: MALE) (BENEFICIARY-OF: SUPPORT-1)
```

1d. The secretary backed her files up.

```
BACKUP-COMPUTER-DATA-1
   (AGENT: SECRETARY-1) (THEME: COMPUTER-FILE-1)
   (TIME:  < FIND-ANCHOR-TIME)
SECRETARY-1
   (AGENT-OF: BACKUP-COMPUTER-DATA-1)
COMPUTER-FILE-1
  (THEME-OF: BACKUP-COMPUTER-DATA-1)
```

1e. The man backed up the car.

```
CHANGE-LOCATION-3
   (CAUSED-BY: HUMAN-4) (THEME: AUTOMOBILE-2)
   (DIRECTION-OF-MOTION: BACKWARD)
   (TIME: < (FIND-ANCHOR-TIME)
HUMAN-4
   (GENDER: MALE) (AGE: > 18)
   (EFFECT: CHANGE-LOCATION-3)
AUTOMOBILE-2
   (THEME-OF: CHANGE-LOCATION-3)
```

1f. The man backed up the driveway.

```
CHANGE-LOCATION-4
   (CAUSED-BY: HUMAN-5) (SPATIAL-PATH: DRIVEWAY-1)
   (DIRECTION-OF-MOTION: BACKWARD)
   (TIME: < FIND-ANCHOR-TIME)
HUMAN-5
   (GENDER: MALE) (AGE: > 18)
   (EFFECT: CHANGE-LOCATION-4)
DRIVEWAY-1
   (SPATIAL-PATH-OF: CHANGE-LOCATION-4)
```

MOTION-EVENT whose DIRECTION-OF-MOTION is BACKWARD. The analyzer compares the meaning of input components with the available lexical entries, scoring potential analyses based on the ontological-semantic conformity of components.

Sentence (1c) (*The woman backed up her husband*). This sentence contains syntactic and semantic ambiguity:

a) *up* could be a particle and the overall meaning could be the idiomatic "support, agree with" (the default reading)
b) *up* could be a particle and the overall meaning could be the idiomatic "cause to move backward"
c) *up* could be a preposition heading the PP *up her husband* and the meaning could be "move (oneself) backwards in a upwards direction" (appropriate if, e.g., the couple were acrobats).

Although all of these readings are possible, most people would prefer the first in the absence of strong contextual priming. The system shows the same preferences since it cannot currently carry out text-level context-based reasoning in open domains. Here we describe how the knowledge sources and processors interact to arrive at the preferred reading.

There are 3 lexical senses of *back* whose semantic structure corresponds to the interpretations in (a) and (b): they all take a subject, the particle *up* and a direct object, and they all permit the particle to come before or after the direct object. The difference between them lies in their semantic structures, which contain the following elements (in simplified format):

back + particle, sense (i)
SUPPORT (AGENT) (BENEFICIARY)

back + particle, sense (ii)
MOTION-EVENT (AGENT) (THEME: VEHICLE)
(DIRECTION: BACKWARDS)

back + particle, sense (iii)
BACKUP-COMPUTER-DATA (AGENT)
(THEME: COMPUTER-FILE)

The analyzer penalizes senses (ii) and (iii) because the head of the direct object, *husband*, is a SPOUSE, and a SPOUSE is neither a type of VEHICLE nor a type of COMPUTER-FILE.[3] Since SPOUSE is a HUMAN and HUMANs can be BENEFICIARIES, sense (i) is selected outright.

The obvious drawback of the current OntoSem analysis strategy is that it excludes the valid, though probably quite rare, reading listed in (b), where the woman is causing her

[3] We actually do not have to list the THEME of BACKUP-COMPUTER-DATA in the lexicon since the analyzer has access to this constraint through the ontological specification of BACKUP-COMPUTER-DATA. AGENTS and BENEFICIARIES are ontologically specified as being, by default, HUMANS.

husband to move backward. There are a number of ways we could allow the system to permit this analysis: e.g., by creating a separate *back* + *particle* sense that takes a HUMAN as the THEME and means "cause someone to be the agent of a motion event", or by introducing a "relaxable-to" semantic constraint on the THEME of the MOTION-EVENT in sense (ii). The reason we do not do this at present is because it will introduce ambiguity that we cannot yet contextually disambiguate, and practicality dictates that such – for the most part, spurious – senses not be included.

Having narrowed down the *back* + *particle* interpretations to just one, the analyzer is still left with one other analysis—the acrobatic one in (c) (the woman crawled up her husband's body). This highly infrequent "productive PP" is not selected due to the analyzer's general preference for encoded lexicalized knowledge over compositional analysis. We have found that the rare miss of a compositional analysis in the presence of a lexicalized one is a small price to pay for avoiding the potential explosion of ambiguity caused if lexicalized readings have no advantage over compositional ones.

Sentence (1d) (*The secretary backed her files up*). Disambiguation of this sentence is readily carried out using the same type of selectional restriction matching as described for (1c). The discontinuous verb + particle is handled by the same lexical sense that handles the other word order: *The secretary backed up her files.*

Sentence (1e) (*The man backed up the car*). This sentence, like (1c), has both a preferred, lexically encoded [verb–particle–direct object] meaning and a more marginal, compositional PP meaning (the man moved back, upwards and over his car, e.g., if a dog were snarling at him). The general preference for lexically encoded meanings over compositional ones excludes the compositional one, and the selectional restriction mapping selects the correct [verb–particle–direct object] sense.

Sentence (1f) (*The man backed up the driveway*). This sentence should be analyzed using the lexical sense of *back* whose syntactic structure requires a PP (**He backed/is backing* is impossible), the head of which is left unspecified to permit the compositional analysis of input like *back up the driveway, down the street, around the house,* etc. The analyzer, in fact, arrives at this analysis unambiguously due to the interaction of two factors. First, a PP is listed as a required component in the target lexical entry, so this sense will receive the same degree of preference as other lexically encoded phrasals (if no PP were listed and the collocation were completely productive, the analysis might have received insufficient preference to be selected over others). Second, the input *driveway,* which maps to DRIVEWAY, does not fit the selectional restrictions on the object in any of the other senses of *back,* so only the desired sense remains.

This Work in the Bigger Picture

We have presented an implemented, semantics-based method of disambiguating different senses of verb–particle constructions, and disambiguating verb–particle constructions from free preposition constructs. Ours is a deep semantics approach that emphasizes the end-to-end analysis of text, from raw input to text-meaning representation. As such, it differs from other extant approaches that, for example, require large annotated corpora (e.g., Gildea and Jurafsky 2002), categorize entities without attaching a semantic interpretation to them (Baldwin and Villavivencio 2002, for verbal particles), attempt disambiguation of only certain entities in a given input (e.g., prepositions, as in O'Hara and Wiebe 2003, Litkowski 2002), or assume the existence of deep knowledge resources that have not actually been developed (e.g., Jensen and Nilsson 2003, who do not specifically discuss the type of disambiguation treated here but do discuss another type of disambiguation of PPs – as in *the treatment of children with diabetes* – using idealized resources).

As implied above, within OntoSem we devote more resources to encoding high-quality knowledge than to developing methods to use lower-quality knowledge, where inadequacies in depth and breadth of coverage of phenomena must then be detected and corrected. However, we are interested in external results that can feed into our acquisition efforts. To give just two examples, the output of Baldwin, Beavers, et al.'s (2003) method for extracting determinerless PPs (e.g., *by train*) from corpora could be useful to guide acquisition of these entities, which would be encoded in the OntoSem lexicon using syntactic patterns with ontologically grounded semantic constraints. Similarly, the work of Villavicencio and Copestake (2003), who sought lexical rules to account for particle interpretations, is also of interest as support for acquisition and as a means of recovering from unexpected (in this case, not yet lexicalized) input.

Zooming Out: Why This is Important

We have just analyzed in some depth a series of examples that were disambiguated correctly by the OntoSem analyzer—an experiment that required no modifications to the OntoSem analyzer and only minimal supplementation to the knowledge resources. That is, the disambiguation of P-Ps is carried out in the same way as the disambiguation of predicate words, arguments, adjuncts, modifiers and all other text input.

There are more meanings of *back* + *up* than the ones we treated here, ranging from the common to the highly marginal. Our point is not to say the last word about *back* + *up*; rather, it is to suggest the benefit of treating prepositions and homographic particles in an integrated environment that allows various types of knowledge to contribute to the heuristics for extracting meaning from text.

Compare this orientation with that reflected in the majority of current work in our field, which tends to treat specific phenomena (like prepositions or particles) in isolation and/or under a variety of assumptions about the future availability of knowledge resources. Such methodologies have undoubtedly led to interesting research programs and innovations, but they can only come to real fruition when incorporated into comprehensive systems that can tie together cutting-edge findings from disparate research paradigms.

When presented with a costly knowledge-rich approach to NLP, it is natural to ask for evidence that the quality of its output exceeds that of cheaper, knowledge-lean methods. Unfortunately, we know of no way to compare these apples and oranges that is both fair and would provide an outcome that is not a given from the outset. That is, stochastic methods are not attempting the depth of semantic analysis that OntoSem is, so they would fail a priori in any evaluation task catered to OntoSem. Conversely, OntoSem, like all knowledge-rich systems, suffers in terms of coverage, so it would show poorly in the kinds of evaluation tasks catered to stochastic methods. (Of course, all evaluation tasks are catered to given approaches; the bias simply tends to be implicit rather than explicit when all competitors work in the same paradigm.)

We are currently pursuing several methodologies to make OntoSem sufficiently robust, in the near term, to do well in the more traditional evaluation tasks that require broad coverage. For example, we have begun experimenting with stochastically supported lexicon and ontology acquisition, which we recognize as one (though not the only!) prerequisite for breaking through the knowledge bottleneck. Concurrently, we are developing practical applications that rely more on precision than recall, like populating a fact repository with semantically analyzed, machine tractable information. Therefore, even before our system reaches the coverage of stochastic systems, its utility is being exploited.

Acknowledgments

Thanks to Tom O'Hara for helpful comments on a draft of this paper.

References

Baldwin, T.; Beavers, J.; van der Beekz, L.; Bond, F.; Flickingery, D.; and Sag, I. 2003. In Search of a Systematic Treatment of Determinerless PPs. In *Proceedings of the ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*, Toulouse, France, pp. 145-56.

Baldwin, T.; and Villavicencio, A. 2002. Extracting the Unextractable: A case study on verb-particles. In *Proceedings of the 6th Conference on Natural Language Learning* (CoNLL-2002), Taipei, Taiwan.

Beale, S.; Nirenburg, S.; and Mahesh, K. 1995. Semantic Analysis in the Mikrokosmos Machine Translation Project. In *Proceedings of Symposium on Natural Language Processing*, Kaset Sart University, Bangkok, Thailand.

Beale, S.; Nirenburg, S.; and McShane, M. 2003. Just-in-Time Grammar. In *Proceedings 2003 International Multiconference in Computer Science and Computer Engineering*. Las Vegas, Nevada

Beale, S.; Lavoie, B.; McShane, M.; Nirenburg, S.; and Korelsky, T. 2004. Question Answering Using Ontological Semantics. In *Proceedings of ACL-2004 Workshop on Text Meaning and Interpretation*. Barcelona, Spain.

Fikes, R.; Frank, G.; and Jenkins, J. 2003. JTP: A System Architecture and Component Library for Hybrid Reasoning. In *Proceedings of the Seventh World Multiconference on Systemics, Cybernetics, and Informatics*. Orlando, Florida.

Gildea, D.; and Jurafsky, D. 2002. Automatic Labeling of Semantic Roles. In *Computational Linguistics* 28(3):245-288.

Huddleston, R.; Pullum, G. 2002. *The Cambridge Grammar of the English Language.* Cambridge University Press.

Jensen, P.A.; and Nillson, J.F. 2003. Ontology-based Semantics for Prepositions. In *Proceedings of the ACL-SIGSEM Workshop: The Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications.* Toulouse, France.

Litkowski, K.C. 2002. Digraph Analysis of Dictionary Preposition Definitions. In *Proceedings of the Association for Computational Linguistics Special Interest Group on the Lexicon*. Philadelphia, PA.

McShane, M.; Beale, S.; and Nirenburg, S. 2004a. Some Meaning Procedures of Ontological Semantics. In *Proceedings of LREC-2004.* Lisbon, Portugal.

McShane, M.; Beale, S.; and Nirenburg, S. 2004b. OntoSem Methods for Processing Semantic Ellipsis. In *Proceedings of the HLT/NAACL 2004 Workshop on Computational Lexical Semantics.* Boston, Mass.

Nirenburg, S.; and Raskin, V. 2004. *Ontological Semantics*. Cambridge, Mass.: The MIT Press.

O'Hara, T.; and Wiebe, J. 2003. Preposition Semantic Classification via Treebank and FrameNet. In *Proceedings of the Conference on Natural Language Learning* (CoNLL-03), Edmonton.

Villavicencio, A.; and Copestake, A. 2003. Verb-Particle Constructions in a Computational Grammar of English. In *Proceedings of the 9th International Conference on HPSG*, Seoul, South Korea.