Précis of *Identification and Representation of Caused Motion Constructions*

Jena D. Hwang
University of Colorado at Boulder

1 Introduction

The root of many challenges in natural language processing (NLP) applications is the fact that humans can convey a single piece of information in numerous and creative ways. Syntactic variations (e.g. *I gave him my book* vs. *I gave my book to him*), the use of synonyms (e.g.*She bought a used car* vs. *She purchased a pre-owned automobile*) and numerous other variations can complicate the semantic analysis and the automatic understanding of a text. Advances in syntactic parsing and semantic role labeling have been a boon to NLP. However, they perform poorly with sentences that do not conform to expected syntax-semantic patterning behavior. Consider the following sentences:

(1) a. The boy threw the ball into the field.

b. The crowd hissed the clown off the stage.

c. The market tilted the economy into recession.

While only the verb in the first sentence is generally identified as a verb of motion that can appear in a caused motion context (i.e. "The boy causes the ball to move"), all three are examples of caused motion constructions (CMCs). The verb *hiss* in sentence 1b is normally considered an intransitive sound emission verb (e.g. *The snake hissed at the clown*), but in this sentence, the verb is coerced into the caused motion interpretation and the semantics of the verb gives the manner in which the movement happened (e.g. *the crowd caused the clown to move off the stage by means of hissing*). The verb *tilt* is a verb of spatial configuration normally taking, as its object, the inclined item (e.g. *He tilted the bottle*). In 1c, the verb is coerced into the caused motion reading and the motion described is abstract rather than physical as in *He tilted the liquid into his mouth and swallowed*. Whether the motion is physical or abstract, the semantics of the sentences is parallel: all three sentences have a causal argument responsible for the event, an argument in motion, and a path that specifies the initial, middle, or final location, state or condition of the argument in motion.

Despite the semantic similarities expressed in these sentences, a sentence is most likely identified as a case of caused motion if the verb includes the semantics of motion in its prototypical sense. Consider the following two sentences, which are first translated into Spanish using Google Translate, then hand translated back into English[1]:

(2) a. They threw him out of the university.
Ellos le echaron fuera de la universidad.
Tr. They threw him out of the university.

b. They hissed him out of the university.
Le silbó fuera de la universidad.
Tr. They whistled to him outside the university.

The verb *throw* in sentence 2a is an example of a verb that most commonly appears in a caused motion usage, while the verb *hiss* in sentence 2b would be considered atypical in this particular semantic context. When the verb *throw* is used, the results are correct. However, the translation for 2b involving the verb *hiss*, which does not include semantics of motion in its prototypical sense, is incorrect and has absolutely no implication that a

[1]Results from Google translate (http://translate.google.com/) run in August 2014.

group of people did something to cause another person to leave the university. Moreover, the resulting Spanish does not at all show the semantic similarities seen in the English sentences.

The reason these coerced usages are challenging for natural language systems is because the current methods for natural language processing rely heavily on the lexical aspects of the language. New and unusual usages of lexical items are problematic, because they are often infrequent in the data and the frequency of their usage is far outnumbered by that of more conventional verbs in the same semantic context. This is especially significant in discriminative learning methods employed by semantic classification tasks, because their predictive capability is dependent on the examples they have seen or experienced during the training stage. If examples for usages like 1b are sparse (i.e., dwarfed in frequency) or lacking in the training data, the classifier cannot gather the necessary statistics about the verb *hiss* used in a caused motion context to be able to correctly predict its meaning when later encountering the usage. It is a well-known problem: various studies have shown that verbs of ambiguous meaning or verbs that do not have training examples contribute to classifier performance drop (Carreras and Marquez, 2005).

Studies have noted that the semantic problem posed by lexical sparsity can be remedied by recognizing the relationship between semantic and syntactic behavior in verbs (Jackendoff, 1990; Levin, 1993). For example, if we can associate *hiss* and *throw* based on their shared syntactic and semantic behavior, such knowledge would help NLP systems with the issue of unseen predicates. The recognition of the importance of classifying verbs according to their syntactic form and their semantic value has led to the creation of the resource VerbNet (Kipper et al., 2006). Following the work by Levin 1993, the verb classes in VerbNet are grouped by their semantic and syntactic similarities, and thereby VerbNet allows the user to abstract away from individual verb types to capture more general features based on the classes' syntactic behavior or semantic types. Indeed, verb classifications based on both meaning and syntactic realization have been shown to help in various tasks such as machine translation, word sense disambiguation, semantic role labeling, acquisition of subcategorization frames, and discourse parsing (Dorr, 1997; Korhonen, 2002; Subba and Eugenio, 2009; Palmer et al., 2009; Merlo and Stevenson, 2002).

Even so, the challenge that VerbNet and all other hand crafted lexically based resources face is that they can never be complete enough to handle all possible usages and combinations that human language can throw at them. For example, VerbNet is capable of providing NLP systems with a way to generalize the semantic and syntactic behavior of a verb in an atypical usage as seen in 1b, if *hiss* is included in the same VerbNet class as the verb *throw*. Thus, the problem is circular: if we want to use VerbNet to correctly interpret the unfamiliar use of the verb *hiss* as used in the sentence, VerbNet needs to have already associated it with the behavior of verbs like *throw*. However, in order for VerbNet to make such an association, the resource creators would already have to be familiar with the use so that *hiss* could be included in the same class as *throw*. Thus, we are left with the question: how do we interpret unseen words and novel lexical usages in the context of supervised learning?

The premise that drives this dissertation, therefore, is that the challenge of generalizability can be addressed by linking semantics more directly with syntax, as theorized by Construction Grammar (Fillmore, 1988; Goldberg, 1995; Kay and Fillmore, 1999; Michaelis, 2004; Goldberg, 2005). Construction Grammar is in agreement with VerbNet's assumption of a relationship between the meaning of word and its syntactic realization, but it goes a step further in proposing that particular syntactic structures can themselves carry meaning, in the same way a lexical item is thought to bear meaning. This theory suggests that the meaning of a sentence arises not only from the lexical items but also from the morpho-syntactic structures or *constructions* the lexical items sit in. In other words, meaning must be interpreted at both the lexical and constructional levels of sentences. This is especially important for sentences in which the lexical semantics of the verb is at odds with the semantics of the sentence.

Thus, if the semantic interpretation is strictly based on the expected semantics of the verb and its arguments, in some cases it will fail to to include the relevant information from the CMC. An accurate semantic role labelling for such sentences requires that NLP classifiers accurately identify these coerced usages in data. Furthermore, once the CMCs are identified and the semantic roles are properly assigned, the sentence would also require an accurate semantic interpretation with appropriate representations that include the semantics of the CMCs.

2 Bigger Picture: VerbNet Constructions

This thesis is part of a greater effort at equipping VerbNet with systematic ways for dealing with coercive usages of verbs (Bonial et al., 2011). Through coercion, a verb is allowed to violate its typical selectional restriction and to be used in atypical syntactic contexts (Pustejovsky and Jezek, 2008; Michaelis, 2005; Goldberg and Jackendoff, 2004)[2]. For example, in sentence 1b, a typically intransitive verb *hiss* (e.g. *The snake hissed at the boy*) is coerced into caused motion context. In the same way, an intransitive action verb like *blink* (e.g. *John blinked at me*) gains a caused motion reading when used with an object of blinking and a path prepositional phrase in *She blinked the snow off her eyelashes*.

VerbNet is currently useful at providing an analysis of the meaning of a sentence that is predictable given the semantics of the verb. However, VerbNet does not have a good analysis for the 'extra' meaning the verb gains through coercion. Consequently, VerbNet's current treatment of the coerced instances is not consistent. Consider the following three sentences:

(3) a. John slouched himself into the chair.

b. The crowd laughed the clown off the stage.

c. Cynthia blinked the snow off her eyelashes.

The verb *slouch* is generally considered a verb of posture along, the verb *laugh* is typically a verb of non-verbal expression, and the verb *blink* is considered a verb of bodily movement. The CMC usages have been handled inconsistently in VerbNet: *slouch* is placed in the RUN-51.3.2 class for verbs of directed motion with expected paths, *laugh* is in the NONVERBAL_EXPRESSION-40.2 class, which arbitrarily includes a path even though path is not fully compatible with the class; and *blink*, in the HICCUP-40.1.1 class, cannot take a path.

We introduce constructional definitions as an alternative to this adhoc and inconsistent treatment. These VerbNet constructions will interact with the current VerbNet classes to project the constructional meaning on to the sentence when the inherent semantics of the verb does not include it. Through these constructions, VerbNet gives a uniform semantic treatment for all sentences of caused motion regardless of the specific lexical meaning of the verb. Figure 1 is a visualization of a CMC instantiated by a verb typical of motion.

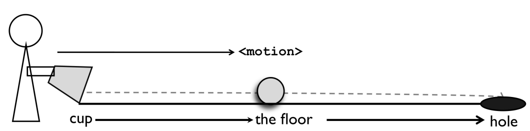


Figure 1: *"Cynthia rolled the ball out of the cup, across the floor, and into a hole."*

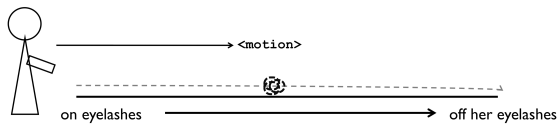
VerbNet would ideally represent the semantics expressed in the above sentence through the following semantic predicates:

```
cause(Cynthia, E)
motion(E, the ball)
rel_to_path(start(E), the ball, the cup)
rel_to_path(during(E), the ball, the floor)
rel_to_path(end(E), the ball, a hole)
```

Cynthia causes an event **E** in which *the ball* is put into motion. At the beginning of **E**, *the ball* is in the cup. During **E**, *the ball* moves across the floor to eventually to be located in *a hole* at the end of the event.

[2]This argument structure level of coercion is also discussed in literature in terms of valence extension or valence augmentation.

Because the verb *roll* and the members of the ROLL-51.3.1 class are verbs of motion, VerbNet would include this syntactic frame within the class. However, the HICCUP-40.1.1 class including the verb *blink* will still not have this frame. Nonetheless, VerbNet will now give these verbs a similar semantic representation when they appear in a caused motion context. Figure 2 shows the visualization of the caused motion event in example 3c and its desirable sentence representation.



cause(*Cynthia*, **E**)
motion(**E**, *the snow*)
body_process(**E**, *Cynthia*)
rel_to_path(start(**E**), *the snow*, *her eyelashes*)
rel_to_path(end(**E**), *the snow*, **not**(*her eyelashes*))

Figure 2: *"Cynthia blinked the snow off her eyelashes."*

Since the class to which the verb *blink* belongs does not a include caused motion frame, the semantics of motion and path information is projected by the semantics of the CMC. More specifically, the `cause()` predicate, the `motion` predicate and the `rel_to_path` predicates is supplied by the semantics of the CMC. One notable difference from a standard VerbNet sentence representation is that the *blink* sentence representation that includes the CMC carries an additional semantic component (i.e. the act of blinking). Thus, what the VerbNet HICCUP-40.1.1 class would then provide is the specialized semantics that are particular to the lexical meaning that instantiates the CMC. In the above example, this lexical information is captured with the `body_process()` predicate from the class HICCUP-40.1.1.

The semantic interaction between CMCs and the ROLL-51.3.1 class or other classes of motion such as THROW-17.1 with members like *kick, lauch*, and *toss* will produce the same results that they would have produced by themselves, since the CMC frame is inherent in the member verbs (e.g. *She carelessly tossed her book on the table*). However, when a CMC is paired with the HICCUP-40.1.1 class, which includes verbs such as *blink*, and *sneeze*, it produces the correct analysis for sentences like *She blinked the snow off her eyelashes* or *Pat sneezed the foam off the cappuccino*.

3 Summary of the Dissertation

Thus, this thesis work focuses on the definition, identification, and, finally, the representation of the CMCs. This work breaks down into four general stages. The first stage is primarily a contribution to theoretical linguistics. In this stage, we expand on the theoretical work from Construction Grammar to develop the semantic types and varieties of CMCs through a detailed corpus study. In the second stage, utilizing the annotation as the training and test data, we use methods in NLP to train automatic CMC classifiers and demonstrate that CMCs can be reliably identified in the corpus data. In the third stage, working within a knowledge representation and reasoning (KRR) framework[3] from Philosophy, we revise the current semantic predicate representation in VerbNet and develop new ones aimed at systematic semantic representation of CMCs. These predicates provide for the representation of CMC sentences, but also give VerbNet a more consistent explicit representation for all verbs that express a path of motion. Finally, by putting all of these three pieces together – linguistics, computational linguistics and knowledge representation and reasoning – we demonstrate that CMC representation can help give the proper semantic representation to coercive sentences even when the verb in the sentence does not include the semantics of CMC.

[3]VerbNet has been already integrated into the KRR ontology CYC (Trumbo, 2006).

3.1 Defining and Annotating Caused Motion Constructions

We carried out a detailed corpus study of CMCs to systematize their defining characteristics into a typological classification using classical linguistic analysis on naturally occurring data, and to adapt the results of the linguistic analysis into a set of annotation guidelines aimed at corpus annotation of CMCs (Hwang et al., 2014). The contribution of this stage of the thesis is in two parts. First, we evaluated the current status of the theoretical work on CMCs as proposed and defined by Goldberg (1995). In a detailed corpus study of over 21K sentences, we found that while Goldberg's semantic constraints serve to account for more prototypical types of CMCs, they fail to account for all of the variety and types of instances seen in data, especially those that are more abstract or metaphorical in meaning. For example, Goldberg's constraints specify that verbs like *encourage* and *persuade* cannot appear in the construction. Nevertheless, there are attested cases of *persuade* and *encourage* appearing in a caused motion-like context (e.g. *Wade's sheer enthusiasm [...] persuaded me into a renewed fascination*). Through the course of this study, we developed three semantic criteria that generalize better over all the instances of CMCs, including both the physical and abstract types. We also examined the metaphorical extensions of CMCs to recognize salient categories that naturally surface in the data.

Second, we established annotation guidelines based on the semantic criteria to aid in the production of consistent and reliable annotation. In an iterative annotation process, the creation of the annotation guidelines and the annotation process run in parallel: the guidelines inform the annotation process and, in turn, the annotation process provides feedback for further refinement of the guidelines. The outcome of this effort is the development of the categorical types of CMCs as they occur in corpus data as seen in the following figure:

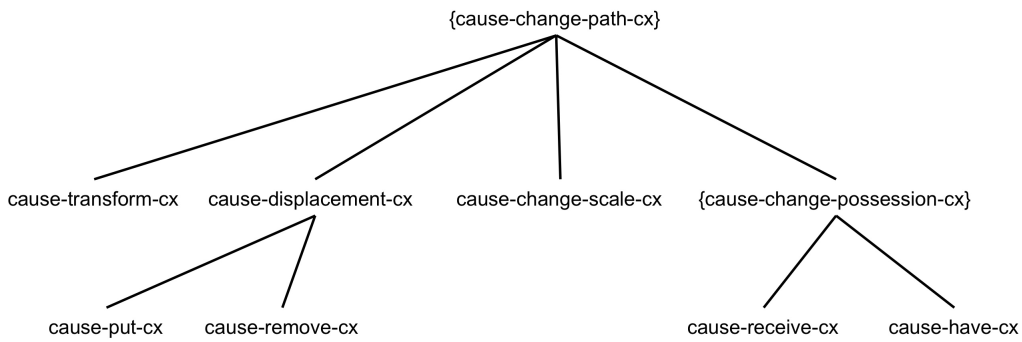


Figure 3: CMC Classification

Our classification of CMCs gives insight into the types and varieties of semantic inferences entailed by CMCs in corpus data. We used categorical labels of CMCs for the annotation of corpus data. Our results show that the introduction of the new semantic criteria and the categorical types contribute to inter-annotator agreement rate, which indicates that they serve to facilitate the annotation process. Moreover, from the standpoint of theoretical linguistics, the successful annotation of the data establishes that CMCs are identifiable in naturally occurring text and that the descriptive linguistic criteria are indeed useful in characterizing CMCs. The annotation provides approximately 21K instances of annotated corpus data for the next stage of this dissertation: the training and testing of automatic CMC classifiers.

3.2 Automatic Identification of Caused Motion Constructions

In this stage of the thesis we take the results of the annotation to build automatic classification of CMCs using supervised methods in machine learning. Our study makes use of SVM (support vector machine) learning model to train an automatic classifier to identify a phenomenon based on exemplars in the training data (Hwang et al., 2010). At the decoding stage (also called testing or evaluation stage), the classifier is asked to predict the label for the phenomenon in a new set of data. The results are then evaluated against the gold standard annotation to gauge the classifiers performance. We utilize the CMC annotation we produced in the previous stage of the dissertation as training and testing data to explore the semantic and syntactic features that would best help characterize CMCs in an automatic learning setting.

Our studies have shown that we can achieve the identification of caused motion instances at the F1 score[4] of 0.818 (using features from Gold Standard parses). This result is at a higher rate than the counterpart inter-annotator agreement F score of 0.667, the best performance that can be realistically expected from an automatic system. We have also shown that semantic information is highly indicative of the caused motion phenomenon, confirming our general intuition that the CMC is a semantic phenomenon. However, we also find that syntax provides scalable features that generalize well across different types of text, producing better results in cross-genre experiments. We have also shown that the downsampling of the negative label has a positive impact on the classification of the labels.

3.3 Representing Caused Motion Constructions in VerbNet

VerbNet currently supplies a large array of semantic predicates for the representation of meaning, making it a natural choice for the representation of CMCs. Additionally, verbs of directed motion (e.g. RUN-51.3.2), verbs of transfer of possession or information (e.g. GIVE-13.1) and verbs of transformation (e.g. Turn-26.6.1), whose semantics directly lend to the description of caused motion events, are currently represented in VerbNet. As we seek to leverage VerbNets semantic predicate representations for CMC representation, we are faced with issues of generalizability: VerbNet had no single set of predicates that could consistently represent the different meaning components of CMCs (Hwang et al., 2013). That is, because VerbNet classes are designed to focus on the protypical or conventional behavior of verbs, semantic coercions that often cross-cut through the established class boundaries are problematic.

In order to improve our ability to use VerbNet and to provide a semantically informed representation of a CMC, we had to augment VerbNet with the information necessary to provide a unified treatment and consistent representation of both the coercive and the conventional usages of verbs primarily by defining a uniform path predicate that could apply to both concrete and physical usages. This stage of the dissertation, effectively affects 123 classes out of the 288 classes currently in existence in VerbNet. More specifically, two new predicates that were formulated through the process replace ten previous predicates that were either semantically inconsistent or too narrow to be used for semantic generalization across multiple classes.

The new semantic predicates established in this study, in conjunction to the existing VerbNet semantic predicates achieve three goals. First, because the semantics of path of motion are not limited to CMCs, the new predicates are also designed to be used for other verbs of motion that include the semantics of path (e.g. intransitive directed motion as in *Eric went from Paris to Lyon* or fictive motion as in *The river meanders through the forest*). Secondly, as discussed in section two, these semantic predicates serve to provide semantic representation for VerbNet's new CMC constructions in such a way that coerced instances such as *Cynthia blinked the snow off her eyelashes* are also appropriately represented in VerbNet, as discussed in more detail in the next stage of the thesis. Thirdly, the improvements in consistency will enhance our ability to use VerbNet in conjunction with KRR ontologies such as CYC or SUMO for knowledge representation and reasoning. The new, improved VerbNet thematic role hierarchy has already been influential in a study of preposition semantic categories for general annotation (Schneider et al., 2015 (under submission).

3.4 CMC Sentence Representations

The fourth stage pulls together the results from the linguistic analysis, the automatic classifier and the semantic representation segments. In this portion of the study, we have produced semantic representation of naturally occurring sentences, with a special focus on coerced instances of CMCs. For the production of sentence representations, we used a simple rule-based implementation system that pulls from the semantic predicates of the

[4] F1 score is the harmonic mean of the precision and the recall values.

CMC definitions and the semantic predicates of the VerbNet class to which the verb is associated. Semantic representations of sentences were produced for a total of 311 sentences that were successfully identified as CMC during the automatic classification stage and identified as cases of coerced instances of CMC.

The results of this study show that by joining the semantics of the CMC to that of the verb's class semantics, we are capable of delivering semantic information that would not be available given a strictly compositional intepretation. Such is the case for the verb *spew*: the SUBSTANCE_EMISSION-43.4 class to which it belongs does not include the CMC in its definitions. Thus, it is by the verb's association to the VerbNet's constructional definition of CMC that we derive the semantic representation for the sentence *The photos showed active geysers* ***spewing*** *sulfurous material into its atmosphere* (i.e. active geysers cause sulfurous material to move into the atmosphere by the means of spewing).

Furthermore, a CMC representation is possible even for those instances whose verb has not been identified with a VerbNet label. For example, the verb *distance* is not included as a member in any of the current VerbNet classes. In spite of this, an appropriate semantic representation is still made available for sentences like *The currency market has begun to distance itself from the volatile stock exchange*: since the sentence has been identified as an instance of CMC, it can receive the necessary semantic predicates directly from the CMC definitions. That is, so long as we can identify a sentence as a CMC, the relevant semantic information will be available via the constructional layer of description.

3.5 Conclusions and interdisciplinary contribution of the dissertation

This dissertation is an interdisciplinary work that pulls together theories and methods in Linguistics, Computer Science, and Philosophy for the identification and representation of CMCs with especial focus on the coerced instances of CMC. Major contributions of this work are as follows. Firstly, we have extended and improved upon the theoretical work of CMCs in Linguistics by carrying out a detailed linguistic analysis over naturally occurring data. Through the process of applying this line of work to the automatic classification in NLP based on machine learning techniques from Computer Science, we confirm and verify that CMCs are identifiable by not only human annotators but also automatic classifiers as well, and that the descriptive linguistic criteria developed in this work are indeed useful in characterizing CMCs. Secondly, this thesis makes a major advance in the knowledge representation and reasoning framework for a major English Lexical Resource, by contributing new VerbNet predicates that significantly improve the consistency of predication across all classes and that make it possible to provide appropriate, explicit representation of the semantics of CMCs. Finally, this thesis demonstrates that CMC representation can help give the proper semantic representation to novel usages of CMCs that coerce the meaning of a different verb type into a CMC reading, even when the verb is not already in a VerbNet class.

The overall contribution of this work is the establishment of the processes involved in identifying and representing constructions in an empirical setting. This work assesses the necessary steps to define and annotate constructions in a corpus setting, train classifiers for constructions, and represent the semantics of constructions through VerbNet predicates. In this study, we provide a framework for implementing similar studies with other constructions and outline the challenges that may be involved in such a process. While we have focused on the identification and representation of CMCs, a similar corpus-driven study can be conducted for other constructions whose sentence representations would not be possible with the semantics of the verb alone, such as adjectival resultatives (e.g. *Mary hammered the metal flat*), ditransitive constructions (e.g. *John lent me a bicycle*), and conative constructions (e.g. *Brian wiped at the counter with a damp rag*).

(Words in text: 3996)

References

Bonial, Claire, Susan Windisch Brown, Jena D. Hwang, Christopher Parisien, Martha Palmer, and Suzanne

Stevenson. 2011. Incorporating coercive constructions into a verb lexicon. In *Proceedings of the ACL 2011 Workshop on Relational Models of Semantics*. Portland, Oregon.

Carreras, Xavier and Lluis Marquez. 2005. Introduction to the conll-2005 shared task: Semantic role labeling. In *Proceedings of CoNLL-2005*.

Dorr, Bonnie. 1997. Large-scale dictionary construction for foreign language tutoring and interlingual machine translation. *Machine Learning* 12(4):271–322.

Fillmore, Charles J. 1988. Regularity and idiomaticity in grammatical constructions: the case of ”let alone”. *Language* 64(3):501–538.

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

Goldberg, Adele E. 2005. *Constructions at Work*. New York: Oxford University Press.

Goldberg, Adele E. and Ray Jackendoff. 2004. The english resultative as a family of constructions. *Language* 80(3):532–568.

Hwang, Jena D., Rodney D. Nielsen, and Martha Palmer. 2010. Towards a domain independent semantics: Enhancing semantic representation with construction grammar. In *Proceedings of the NAACL HLT Workshop on Extracting and Using Constructions in Computational Linguistics*, pages 1–8. Los Angeles, California: Association for Computational Linguistics.

Hwang, Jena D., Martha Palmer, and Annie Zaenen. 2013. Representing paths of motion in representing paths of motion in VerbNet. In T. H. King and V. de Paiva, eds., *From Quirky Case to Representing Space*. CSLI Online Publications.

Hwang, Jena D., Annie Zaenen, and Martha Palmer. 2014. Criteria for identifying and annotating caused motion constructions in corpus data. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Reykjavik, Iceland.

Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, Massachusetts: The MIT Press.

Kay, Paul and Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: the what's x doing y? construction. *Language* .

Kipper, Karin, Anna Korhonen, Neville Ryant, and Martha Palmer. 2006. A large-scale extension of VerbNet with novel verb classes. In *Proceedings of EURALEX*. Turin, Italy.

Korhonen, Anna. 2002. Semantically motivated subcategorization acquisition. In *Proceedings of the ACL Workshop on Unsupervised Lexical Acquisition*, pages 51–58. Philadelphia, PA.

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

Merlo, Paola and Suzanne Stevenson. 2002. Structure and frequency of lexical semantic classes: Consequences for learning. In *Workshop on Quantitative Investigations in Theoretical Linguistics (QITL)*. Osnabruek, Germany.

Michaelis, Laura A. 2004. Type shifting in construction grammar: An integrated approach to aspectual coercion. *Cognitive Linguistics* 15:1–67.

Michaelis, Laura A. 2005. Entity and event coercion in a symbolic theory of syntax. In *Construction Grammar(s): Cognitive Grounding and Theoretical Extensions*, vol. 3 of *Constructional Approaches to Language*, pages 45–87. Amsterdam: Benjamins.

Palmer, Martha, Jena D. Hwang, Susan Windisch Brown, Karin Kipper Schuler, and Arrick Lanfranchi. 2009. Leveraging lexical resources for the detection of event relations. In *Proceedings of the AAAI 2009 Spring Symposium on Learning by Reading*. Stanford, CA.

Pustejovsky, James and Elisabetta Jezek. 2008. Semantic coercion in language: Beyond distributional analysis. *Italian Journal of Linguistics* .

Schneider, Nathan, Jena D. Hwang, Vivek Srikumar, and Martha Palmer. 2015 (under submission). Paths in a semantic role hierarchy for prepositions. In *Proceedings 13th International Cognitive Linguistics Conference*. Newcastle, England.

Subba, Rajen and Barbara Di Eugenio. 2009. An effective discourse parser that uses rich linguistic information. In *Proceedings of HLT/NAACL-2009*, pages 566–574.

Trumbo, Derek. 2006. *Improving Lexical Resources*. Master's thesis, University of Colorado at Boulder.