

The Semantic Proto-Role Linking Model

Aaron Steven White
Science of Learning Institute
Johns Hopkins University
aswhite@jhu.edu

Kyle Rawlins
Cognitive Science
Johns Hopkins University
kgr@jhu.edu

Benjamin Van Durme
Computer Science
Johns Hopkins University
vandurme@cs.jhu.edu

Abstract

We propose the *semantic proto-role linking model*, which jointly induces both predicate-specific semantic roles and predicate-general *semantic proto-roles* based on semantic proto-role property likelihood judgments. We use this model to empirically evaluate Dowty’s thematic proto-role linking theory.

1 Introduction

A *linking theory* explains how predicates’ *semantic arguments*—e.g. `HITTER`, `HITTEE`, and `HITTING-INSTRUMENT` for *hit*—are mapped to their *syntactic arguments*—e.g. subject, direct object, or prepositional object (see Fillmore 1970; Zwicky 1971; Jackendoff 1972; Carter 1976; Pinker 1989; Grimshaw 1990; Levin 1993).

- (1) a. [John]_{HITTER} hit [the fence]_{HITTEE}.
b. [The stick]_{INST} hit [the fence]_{HITTEE}.

A *semantic role labeling* (SRL) system implements the *inverse* of a linking theory: where a linking theory maps a predicate’s observed semantic arguments to its latent syntactic arguments, an SRL system maps a predicate’s observed syntactic arguments to its latent semantic arguments (see Gildea and Jurafsky 2002; Litkowski 2004; Carreras and Marquez 2004; Marquez et al. 2008).

SRL is generally treated as a supervised task—requiring semantic role annotation, which is expensive, time-consuming, and hard to scale. This has led to the development of unsupervised systems for *semantic role induction* (SRI), which induce predicate-specific roles—cf. PropBank roles (Palmer et al., 2005)—from syntactic and lexical features of a predicate and its arguments.

One approach to SRI that has proven fruitful is to explicitly implement linking as a compo-

nent of generative (cf. Grenager and Manning, 2006) or discriminative (cf. Lang and Lapata, 2010) models. But while most SRI systems have some method for generalizing across predicate-specific roles, few explicitly induce predicate-general roles—cf. VerbNet roles (Kipper-Schuler, 2005)—separately from predicate-specific roles. This is a missed opportunity, since the nature of such roles is a contentious topic in the theoretical literature, and the SRI task seems likely to be useful for approaching questions about them in an empirically rigorous way.

We focus in particular on empirically assessing the semantic proto-role theory developed by Dowty (1991). We propose the *semantic proto-role linking model* (SPROLIM), which jointly induces both predicate-specific roles and predicate-general *semantic proto-roles* (Dowty, 1991) based on semantic proto-role property likelihood judgments (Reisinger et al., 2015; White et al., 2016).

We apply SPROLIM to Reisinger et al.’s proto-role property annotations of PropBank. To evaluate SPROLIM’s ability to recover predicate-specific roles, we compare the predicate-specific roles it induces against PropBank, finding that SPROLIM outperforms baselines that do not distinguish predicate-specific and predicate-general roles. We then compare the predicate-general roles that SPROLIM induces against those Dowty proposes, finding a predicate-general role that matches Dowty’s `PROTOAGENT`. Finally, our work could be viewed as an approach to associating a vector-space semantics to the categorical labels of existing type-level semantic role resources, and so we release a resource that maps from PropBank roles to semantic vectors as fit by SPROLIM.

2 Related work

Prior work in SRI has tended to focus on using syntactic and lexical features to cluster arguments

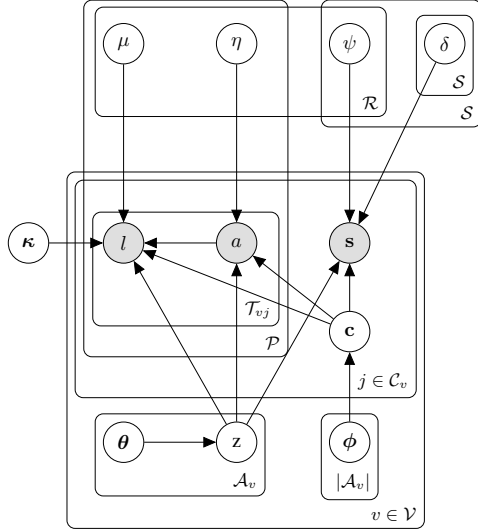


Figure 1: Plate diagram for SPROLIM

into semantic roles. Swier and Stevenson (2004) introduce the first such system, which uses a bootstrapping procedure to first associate verb tokens with frames containing typed slots (drawn from VerbNet), then iteratively compute probabilities based on cooccurrence counts and fill unfilled slots based on these probabilities.

Grenager and Manning (2006) introduce the idea of generating syntactic position based on a latent semantic role representation learned from syntactic and selectional features. Lang and Lapata (2010) expand on Grenager and Manning (2006) by introducing the notion of a *canonicalized linking*. The idea behind canonicalization is to account for the fact that the syntactic argument that a particular semantic argument is mapped to can change depending on the syntax. For instance, when *hit* is passivized, the HITTEE argument is mapped to subject position, where it would normally be mapped to object position.

(2) [The fence]_{HITTEE} was hit.

We incorporate both ideas into our Semantic Proto-Role Linking Model (SPROLIM).

SRI approaches that do not explicitly incorporate the idea of a linking theory have also been popular. Lang and Lapata (2011a, 2014) use graph clustering methods and Lang and Lapata (2011b) use a split-merge algorithm to cluster arguments based on syntactic context. Titov and Klementiev (2011) use a non-parametric clustering method based on the Pitman-Yor Process, and Titov and Klementiev (2012) propose nonparametric cluster-

Algorithm 1 Semantic Proto-Role Linking Model

```

1: for verb type  $v \in \mathcal{V}$  do
2:   for argument type  $i \in \mathcal{A}_v$  do
3:     draw semantic protorole  $z_{vi} \sim \text{Cat}(\theta_{vi})$ 
4:   for verb token  $j \in \mathcal{C}_v$  do
5:     draw canonicalization  $k \sim \text{Cat}(\phi_v | \mathcal{T}_{vj})$ 
6:      $c_{vj} \leftarrow$  element of symmetric group  $S_{|\mathcal{T}_{vj}|, k}$ 
7:     let  $\mathbf{r} : |\mathcal{T}_{vj}|$ -length tuple
8:     for argument token  $t \in \mathcal{T}_{vj}$  do
9:        $r_t \leftarrow$  semantic protorole  $z_{vc_{vj}t}$ 
10:      for property  $p \in \mathcal{P}$  do
11:        draw  $a_{vjt} \sim \text{Bern}(\eta_{r_{vjt}p})$ 
12:        if  $a_{vjt} = 1$  then
13:          draw  $l_{vjt} \sim \text{Cat}(\text{Ord}_\kappa(\mu_{r_{t}p}))$ 
14:      let  $\rho : |\mathcal{S}^{|\mathcal{T}_{vj}|}|$ -length vector
15:      for linking  $s' \in \mathcal{S}^{|\mathcal{T}_{vj}|}$  do
16:         $\rho_{s'} \leftarrow \prod_t \text{softmax}(\psi_{r_t} + \sum_{o \neq t} \delta_{s'_t s'_o})$ 
17:      draw linking  $k \sim \text{Cat}(\rho)$ 
18:       $s_{vj} \leftarrow S_k^{|\mathcal{T}_{vj}|}$ 

```

ing models based on the Chinese Restaurant Process (CRP) and distance dependent CRP.

While each of these SRI systems have some method for generalizing across predicate-specific roles, few induce explicit predicate-general roles, like AGENT and PATIENT, separately from predicate-specific roles. One obstacle is that there is no agreed upon set of roles in the theoretical literature, making empirical evaluation difficult. One reason that such a set does not exist is that reasonably wide-coverage linking theories require an ever-growing number of roles to capture linking regularities—a problem that Dowty (1991) refers to as *role fragmentation* (see also Dowty, 1989).

As a solution to role fragmentation, Dowty proposes the *proto-role linking theory* (PRLT). Instead of relying on categorical roles, such as AGENT and PATIENT—like traditional linking theories do—PRLT employs a small set of relational properties (e.g. *volition*, *instigation*, *change of state*, etc.) that a predicate can entail about its arguments. Dowty partitions these relational properties into two sets, indexed by two *proto-roles*: PROTOAGENT and PROTOPATIENT. The syntactic position that a particular predicate-specific role is mapped to is then determined by how many properties from each set hold of arguments that fill that role. The reason PROTOAGENT and PROTOPATIENT are known as *proto-roles* is that they amount to role *prototypes* (Rosch and Mervis, 1975): a particular predicate-specific role can be closer or further from a PROTOAGENT or PROTOPATIENT depending on its properties.

Reisinger et al. (2015) crowd-sourced annota-

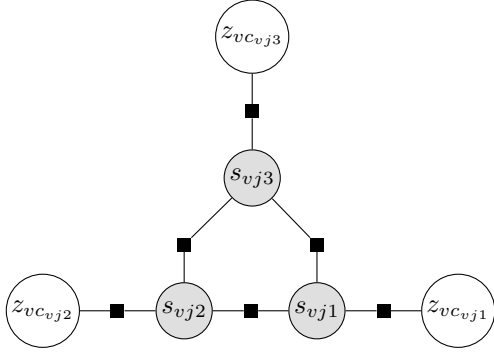


Figure 2: Linking model factor graph for token j of predicate v with three arguments.

tions of Dowty’s proto-role properties by gathering answers to simple questions about how likely, on a five-point scale, it is that particular relational properties hold of arguments in PropBank (cf. Kako, 2006; Greene and Resnik, 2009; Hartshorne et al., 2013). We use these annotations, known as SPR1 (White et al., 2016), to train our *semantic proto-role linking model* (SPROLIM).¹

3 Semantic Proto-Role Linking Model

SPROLIM implements a generalization of Dowty’s semantic proto-role linking theory that allows for any number of proto-roles—i.e. predicate-general roles. Figure 1 shows a plate diagram for the full model, and Algorithm 1 gives its generative story. There are two main components of SPROLIM: (i) the *property model* and (ii) the *mapping model*.

Property model The property model relates each predicate-general role—i.e. proto-role—to (i) the likelihood that a property is applicable to an argument with that role and, (ii) if applicable, how likely it is the property holds of that argument.

We implement this model using a cumulative link logit hurdle model (see Agresti, 2014). In this model, each semantic proto-role $r \in \mathcal{R}$ is associated with two $|\mathcal{P}|$ -length real-valued vectors: η_r , which gives the probability that each property p is applicable to an argument that has role r , and μ_r , which corresponds to the likelihood of each property $p \in \mathcal{P}$ when an argument has role r .

In the hurdle portion of the model, a Bernoulli probability mass function for applicability $a \in \{0, 1\}$ is given by $\mathbb{P}(a | \eta) = \eta^a (1 - \eta)^{1-a}$. What makes this a hurdle model is that the rating probability only kicks in if the rating crosses the applicability “hurdle” (cf. Mullahy, 1986). The pro-

cedural way of thinking about this is that, first, a rater decides whether a property is applicable; if it is not, they stop; if it is, they generate a rating. The joint probability of l and a is then defined as

$$\mathbb{P}(l, a | \mu, \eta, \kappa) \propto \mathbb{P}(a | \eta) \mathbb{P}(l | \mu, \kappa)^a$$

In the cumulative link logit portion of the model, a categorical probability mass function with support on the property likelihood ratings $l \in \{1, \dots, 5\}$ is determined by a latent μ and a nondecreasing real-valued cutpoint vector κ .

$$\mathbb{P}(l = j | \mu, \kappa) = \begin{cases} 1 - q_{j-1} & \text{if } j = 5 \\ q_j - q_{j-1} & \text{otherwise} \end{cases}$$

where $q_j \equiv \text{logit}^{-1}(\kappa_{j+1} - \mu)$ and $q_0 \equiv 0$. In Algorithm 1, we denote the parameters of this distribution as $\text{Ord}_{\kappa}(\mu)$.

Mapping model The mapping model has two components: (i) the canonicalizer, which maps from argument tokens to predicate-specific roles, and (ii) the linking model, which maps from predicate-specific roles to syntactic positions.

We implement the canonicalizer by assuming that, for each predicate (verb) v , there is some canonical ordering of its predicate-specific roles and that for each sentence (clause) $j \in \mathcal{C}_v$ that v occurs in, there is some permutation of v ’s argument tokens in that sentence that aligns them with their predicate-specific role in the canonical order. Denoting the set of argument tokens in sentence j with \mathcal{T}_{vj} , the set of possible mappings is the symmetric group $S_{|\mathcal{T}_{vj}|}$. We place a categorical distribution with parameter ϕ_v on this group.

We implement the linking model using the conditional random field whose factor graph is depicted in Figure 2. This diagram corresponds to the s node and all of its parents in Figure 1.

4 Experiments

In this experiment, we fit SPROLIM to the SPR1 data and investigate the predicate-specific and predicate-general roles it learns.²

Baseline models We use two kinds of Gaussian Mixture Models (GMMs) as baselines: one that uses only the property judgments associated with each argument and another that uses both

¹SPR1 is available at <http://decomp.net>.

²All code, along with the learned predicate and role representations, are available at <http://decomp.net>.

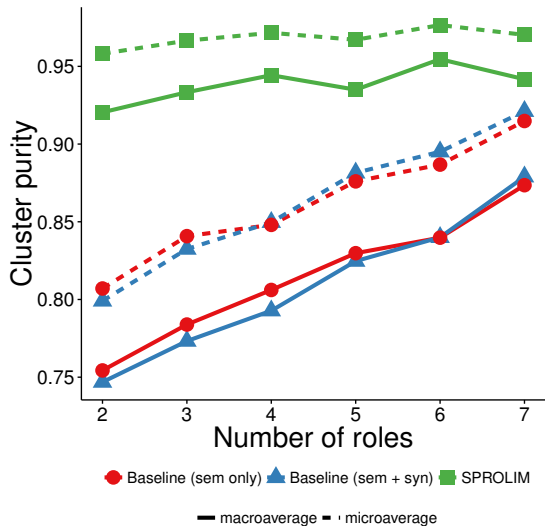


Figure 3: Cluster purity for predicate-specific roles with baselines and SPROLIM.

those property judgments and the syntactic position. We treat each GMM component as a semantic role, extracting each argument’s role by taking the maximum over that argument’s mixture distribution. Since there is no principled distinctions among GMM components, these baselines implement systems that does not distinguish between predicate-specific and predicate-general roles.

Model fitting To fit SPROLIM, we use projected gradient descent with AdaGrad (Duchi et al., 2011) to find an approximation to the maximum likelihood estimates for Θ , Φ , \mathbf{M} , \mathbf{E} , Ψ , Δ , and κ , with the categorical variables \mathbf{Z} and \mathbf{C} integrated out of the likelihood. To fit the GMM baselines, we use Expectation Maximization.

Results Following Lang and Lapata (2010) and others, we evaluate the model using cluster purity.

$$\text{purity}(C, T) = \sum_i^{|C|} \frac{1}{|c_i|} \max_j |c_i \cap t_j|$$

where $C = \{c_i\}$ is the partition of a predicate’s arguments given by a model, and $T = \{t_j\}$ is some ground truth partition—here, PropBank roles.

Figure 3 shows the micro- and macro-average cluster purity for both the GMM baselines and SPROLIM fit with differing numbers of semantic roles. We see that even with only two predicate-general proto-roles, SPROLIM is better able to assign correct predicate-specific roles than the two baseline GMMs. SPROLIM reaches maximum cluster purity at six proto-roles.

Figure 4 shows the estimates of the property likelihood centroids \mathbf{L} for $|\mathcal{R}| \in \{2, 6\}$. Columns give the prototype centroid for a single proto-role.

At $|\mathcal{R}| = 2$, the first proto-role centroid corresponds nearly perfectly to the PROTOAGENT role proposed by Dowty. Furthermore, by inspecting the role-syntax associations Ψ , we see that this proto-role is more strongly associated with the subject position than proto-role 2, and so we henceforth refer to it as the PROTOAGENT role.

A proto-role analogous to the PROTOAGENT role is found for all other values of $|\mathcal{R}|$ that we fit. For instance, at $|\mathcal{R}| = 6$, the first proto-role centroid is highly correlated with the first proto-role centroid at $|\mathcal{R}| = 2$. The only difference between this centroid and the one found at $|\mathcal{R}| = 2$ is that the one at $|\mathcal{R}| = 6$ loads even more positively on Dowty’s proto-agent properties.

At $|\mathcal{R}| = 6$, the second proto-role centroid appears to be a modified version of the PROTOAGENT role that does not require physical existence or sentience and is negatively associated with physical contact. By investigating the proto-role mixtures Θ for each argument, we see that this captures cases of nonsentient or abstract—but still agentive—subjects—e.g. *Mobil* in (3).

- (3) *Mobil restructured* the entire company during an industrywide shakeout.

The rest of the roles are more varied. For $|\mathcal{R}| = 2$, the second proto-role centroid loads negatively (or near zero) on all PROTOAGENT properties, and really, all other properties besides MANIPULATED BY ANOTHER. This non-PROTOAGENT role appears to split into four separate roles at $|\mathcal{R}| = 6$, three of which load heavily on *manipulated by another* (proto-roles 4-6) and the fourth of which (proto-role 3) requires *makes physical contact*. Each of these four non-PROTOAGENT roles might be considered to be different flavors of PROTOPATIENT, which does not appear to be a unified concept. This is corroborated by examples of arguments that load on each of these four proto-roles.

For instance, the objects of *sign*, *want*, and *divert* load heavily on the third proto-role.

- (4) a. President Bush **signed** a disaster declaration covering seven CA counties.
 b. The U.S. **wants** a higher won to make South Korea’s exports more expensive and help trim Seoul’s trade surplus.

- c. They **divert** *law-enforcement resources* at a time they are most needed for protecting lives and property.

The subjects of verbs like *date*, *stem*, and *recover* (in their intransitive form) load heavily on the fourth proto-role.

- (5) a. *His interest in the natural environment* **dates** from his youth.
 b. *Most of the telephone problems* **stemmed** from congestion.
 c. *Junk bonds* also **recovered** somewhat, though trading remained stalled.

The objects of verbs like *reduce*, *lower*, and *slash* load heavily on the fifth proto-role.

- (6) a. The firm **reduced** *those stock holdings* to about 70%.
 b. It also **lowered** *some air fares*.
 c. Robertson Stephens **slashed** *the value of the offering* by 7%.

And the objects of verbs like *gain*, *lose*, and *drop*, which tend to involve measurements, load heavily on the sixth proto-role.

- (7) a. Fujisawa **gained** 50 to 2,060.
 b. A&W Brands **lost** 1/4 to 27 .
 c. B.F. Goodrich **dropped** 1 3/8 to 49 1/8 .

This last category is interesting because it raises a question about how sensitive SPROLIM is to the particular domain on which the proto-role properties are annotated. For instance, outside of newswire, the senses of the verbs in (7) are less likely to include measure arguments, and so perhaps SPROLIM would not find such a proto-role in annotations of text from a different genre.

We believe this warrants further investigation. But we also note that (7) does not exhaust the kinds of arguments that load heavily on the sixth proto-role: the objects of *consume* and *borrow* (among many others) also do so.

- (8) a. In fact, few **consume** *much of anything*.
 b. All they are trying to do is **borrow** *some of the legitimacy of the Bill of Rights*.

The fact that the arguments in (8) are at least superficially unlike the measure arguments found in (7) may suggest that SPROLIM is discovering that measure arguments such as those in (7) fall into a

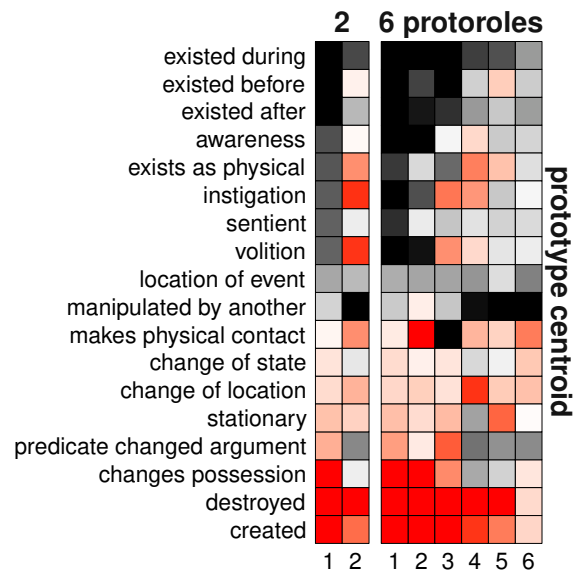


Figure 4: Heatmap of prototype centroids for property likelihood ratings for models with 2 proto-roles and 6 proto-roles. Black is + and red is -.

larger category, in spite of genre-related biases.

5 Conclusion

In this paper, we proposed the *semantic proto-role linking model*, which jointly induces both predicate-specific semantic roles and predicate-general *semantic proto-roles* based on semantic proto-role property likelihood judgments. We used this model to empirically evaluate Dowty’s thematic proto-role linking theory, confirming the existence of Dowty’s PROTOAGENT role but finding evidence that his PROTOPATIENT role may consist of at least four subtypes.

We have three aims for future work: (i) to assess how robust the proto-roles we induce here are to genre effects; (ii) to assess whether languages differ in the set of proto-roles they utilize; and (iii) to extend this model to incorporate annotations that semantically decompose noun meanings and verb meanings in theoretically motivated ways (cf. White et al., 2016).

Acknowledgments

This work was supported in part by the JHU HLT-COE and DARPA LORELEI. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes. The views and conclusions contained in this publication are those of the authors and should not be interpreted as representing official policies or endorsements of DARPA or the U.S. Government.

References

- Alan Agresti. *Categorical Data Analysis*. John Wiley & Sons, 2014. ISBN 1-118-71085-1.
- Xavier Carreras and Llus Marquez. Introduction to the CoNLL-2004 shared task: Semantic role labeling. In *Proceedings of the Ninth Conference on Computational Natural Language Learning*, pages 152–164. Association for Computational Linguistics, 2004.
- Richard Carter. Some linking regularities. In *On Linking: Papers by Richard Carter*, Lexicon Project Working Papers (Vol. 25). MIT Center for Cognitive Science, Cambridge, MA, 1976.
- David Dowty. Thematic proto-roles and argument selection. *Language*, 67(3):547–619, 1991.
- David R. Dowty. On the semantic content of the notion of thematic role. In *Properties, types and meaning*, pages 69–129. Springer, 1989.
- John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *The Journal of Machine Learning Research*, 12:2121–2159, 2011.
- Charles John Fillmore. The grammar of hitting and breaking. In R.A. Jacobs and P.S. Rosenbaum, editors, *Readings in English Transformational Grammar*, pages 120–133. Ginn, Waltham, MA, 1970.
- Daniel Gildea and Daniel Jurafsky. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3):245–288, 2002.
- Stephan Greene and Philip Resnik. More than words: Syntactic packaging and implicit sentiment. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 503–511. Association for Computational Linguistics, 2009. ISBN 1-932432-41-8.
- Trond Grenager and Christopher D. Manning. Unsupervised discovery of a statistical verb lexicon. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pages 1–8. Association for Computational Linguistics, 2006. ISBN 1-932432-73-6.
- Jane Grimshaw. *Argument structure*. MIT Press, Cambridge, MA, 1990. ISBN 0262071258.
- Joshua K. Hartshorne, Claire Bonial, and Martha Palmer. The VerbCorner Project: Toward an Empirically-Based Semantic Decomposition of Verbs. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1438–1442, 2013.
- Ray Jackendoff. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, MA, 1972. ISBN 0-262-10013-4.
- Edward Kako. Thematic role properties of subjects and objects. *Cognition*, 101(1):1–42, 2006.
- Karin Kipper-Schuler. *VerbNet: A broad-coverage, comprehensive verb lexicon*. PhD thesis, University of Pennsylvania, 2005.
- Joel Lang and Mirella Lapata. Unsupervised induction of semantic roles. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 939–947. Association for Computational Linguistics, 2010. ISBN 1-932432-65-5.
- Joel Lang and Mirella Lapata. Unsupervised semantic role induction with graph partitioning. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1320–1331. Association for Computational Linguistics, 2011a. ISBN 1-937284-11-5.
- Joel Lang and Mirella Lapata. Unsupervised semantic role induction via split-merge clustering. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, volume 1, pages 1117–1126. Association for Computational Linguistics, 2011b. ISBN 1-932432-87-6.
- Joel Lang and Mirella Lapata. Similarity-driven semantic role induction via graph partitioning. *Computational Linguistics*, 40(3):633–669, 2014.
- Beth Levin. *English verb classes and alternations: A preliminary investigation*. University of Chicago Press, 1993. ISBN 0226475336.
- Ken Litkowski. Senseval-3 task: Automatic labeling of semantic roles. *Senseval-3: Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text*, 1:141–146, 2004.

- Lluís Marquez, Xavier Carreras, Kenneth C. Litkowski, and Suzanne Stevenson. Semantic role labeling: an introduction to the special issue. *Computational linguistics*, 34(2):145–159, 2008.
- John Mullahy. Specification and testing of some modified count data models. *Journal of econometrics*, 33(3):341–365, 1986.
- Martha Palmer, Daniel Gildea, and Paul Kingsbury. The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106, 2005.
- Steven Pinker. *Learnability and Cognition: The Acquisition of Argument Structure*. MIT Press, Cambridge, MA, 1989. ISBN 0-262-51840-6.
- Drew Reisinger, Rachel Rudinger, Francis Ferraro, Craig Harman, Kyle Rawlins, and Benjamin Van Durme. Semantic Proto-Roles. *Transactions of the Association for Computational Linguistics*, 3:475–488, 2015.
- Eleanor Rosch and Carolyn B. Mervis. Family resemblances: Studies in the internal structure of categories. *Cognitive psychology*, 7(4):573–605, 1975.
- Robert S. Swier and Suzanne Stevenson. Unsupervised semantic role labelling. In *Proceedings of EMNLP*, volume 95, page 102, 2004.
- Ivan Titov and Alexandre Klementiev. A Bayesian model for unsupervised semantic parsing. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, volume 1, pages 1445–1455. Association for Computational Linguistics, 2011. ISBN 1-932432-87-6.
- Ivan Titov and Alexandre Klementiev. Crosslingual induction of semantic roles. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers*, volume 1, pages 647–656. Association for Computational Linguistics, 2012.
- Aaron Steven White, Drew Reisinger, Keisuke Sakaguchi, Tim Vieira, Sheng Zhang, Rachel Rudinger, Kyle Rawlins, and Benjamin Van Durme. Universal decompositional semantics on universal dependencies. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1713–1723, Austin, TX, 2016. Association for Computational Linguistics.
- Arnold M. Zwicky. In a manner of speaking. *Linguistic Inquiry*, 2(2):223–233, 1971.