

A voting scheme to detect semantic underspecification

Héctor Martínez Alonso, Bolette Sandford Pedersen, Núria Bel

University of Copenhagen, Universitat Pompeu Fabra
Copenhagen (Denmark), Barcelona (Spain)
alonso@hum.ku.dku

Abstract

The following work describes a voting system to automatically classify the sense selection of the complex types Location/Organization and Container/Content, which depend on regular polysemy, as described by the Generative Lexicon (Pustejovsky, 1995). This kind of sense alternations very often presents semantic underspecification between its two possible selected senses. This kind of underspecification is not traditionally contemplated in word sense disambiguation systems, as disambiguation systems are still coping with the need of a representation and recognition of underspecification (Pustejovsky, 2009). The data are characterized by the morphosyntactic and lexical environment of the headwords and provided as input for a classifier. The baseline decision tree classifier is compared against an eight-member voting scheme obtained from variants of the training data generated by modifications on the class representation and from two different classification algorithms, namely decision trees and k-nearest neighbors. The voting system improves the accuracy for the non-underspecified senses, but the underspecified sense remains difficult to identify.

Keywords: voting system, regular polysemy, semantic underspecification

1. Introduction

Semantic underspecification has been object of interest in the last years, from the Generative Lexicon (GL) framework for lexical semantics (Pustejovsky 1995) to GL-compliant lexicon and ontology building (Lenci et al, 2000; Buitelaar 1998), as well as other GL-based sense-disambiguation works, such as Rumshisky et al (2007). Disambiguation systems, however, are still coping with the need of a representation and recognition of underspecification (Pustejovsky, 2009).

The cause for a lexical sense to be underspecified is very often regular polysemy. Regular polysemy is the phenomenon whereby a word that belongs to a semantic type can act as a member of another semantic type without incurring in metaphor, as this change of type is the result of metonymy (Apresjan 1974, Pustejovsky 1995).

In spite of the GL's computational aim, implementations that examine the actual computational feasibility of the GL are few. Moreover there is no overt attempt to identify the possible three behaviors of a dot type, as the dot predication has not been computationally tackled, given the lack of strategies to capture meaning underspecification. The goal of the experiment is to use classification to identify the cases of dot predications in two datasets which correspond to two different dot types: Location/Organization and Container/Content.

2. Regular polysemy

Very often a word that belongs to a semantic type, like Location, can behave as a member of another semantic type, like Organization, as shown by the following examples from the American National Corpus or ANC (Ide and Macleod, 2001):

- a) *Manuel died in exile in 1932 in England.*
- b) *England was being kept busy with other concerns*

- c) *England was, after all, an important wine market*

In case a), *England* refers to the English territory (Location), whereas in b) it refers to England as a political entity (Organization). The third case refers to both the English territory and the English people. The ability of certain words to switch between semantic types in a predictable manner is named by different authors as logical, complementary, systematic or regular polysemy.

Apresjan (1974) offers the following definition: "*For any word that has a meaning of type 'A', is true that it can be used in a meaning of type 'B' as well [...] Regular polysemy is triggered by metonymy, whereas irregular polysemy is triggered by other metaphorical processes.*"

Pustejovsky (1995) provides a different wording: "*I will define logical polysemy as a complementary ambiguity where there is no change of lexical category, and the multiple senses of the word have overlapping, dependent or shared meanings.*"

Pustejovsky's definition focuses on the relations between the senses (and the possibility of overlap between them) in the cases of regular polysemy, whereas Apresjan focuses on the fact that the alternation of senses is caused by metonymy. From these two definitions we understand regular polysemy as a phenomenon whereby a word that belongs to a semantic type can act as a member of another semantic type without incurring in metaphor, as this change of type is the result of metonymy. Some well-known examples are:

- d) Container for content: *He drank a whole glass.*
- e) Location for organization: *France elects a new president.*
- f) Property for subject of property: *The authorities arrived quickly.*
- g) Producer for product: *I drive a Honda.*

h) Artifact for information: *It is a sad book.*

In order to be able postulate that words can shift between semantic types we first need to define a list of semantic types, therefore our understanding of regular polysemy is ontology-dependent. The majority of ontologies would portray a stone as a Physical Object and would also provide a Substance sense. A book is just a Document type for some ontologies (Jezek and Quochi, 2010) while for other frameworks a book is something that is considered an Artifact as well as Information, while it is still regarded as fundamentally monosemous because its two class-based senses are not contrastive (Pustejovsky, 1995) and are metonymically related (cf. 2.1).

More importantly, the two metonymically related senses of *book* can be selected simultaneously in a zeugmatic sentence without being anomalous, something that does not happen with for instance *crane*. This phenomenon extends to all cases of regular polysemy.

i) *This book is difficult both to read and carry around*

j) *? Cranes live in swamps and are used in construction*

We consider that *book* is a word that belongs to a complex semantic type (Artifact/Information) and presents a form of regular polysemy whereas *crane* is a word that presents irregular polysemy and does not belong to a complex type, because its senses are contrastive—i.e. mutually exclusive—and cannot co-occur.

2.1. The dot type

The previous section describes the differences between irregular and regular polysemy, and how the latter manifests in non-mutually exclusive senses. This section describes the dot type and how it relates to the aforementioned properties of regular polysemy.

The Generative Lexicon or GL (Pustejovsky, 1995) is a theoretical framework of lexical semantics that tackles the description of the generativity of word meaning. The GL makes use of a series of theoretical objects like *qualia structure*, *type coercion* and *dot type*.

The dot object or dot type is the GL term to give account for words that are inherently members of two non-vertically related semantic classes in an ontology, that is, two semantic classes in which no class subsumes the other: “*The dot object is the logical pairing of the senses denoted by the individual types in the complex type. That is [...] we assume that nominals such as book are a sort of container which are further specified as a relation between a physical object and the textual information contained within it.[...]The dot object is a type which necessarily incorporated the meanings (i.e. types) of its simple types into the complex object.*” (Pustejovsky, 1995).

Incorporating a new concept into a theoretical framework needs a rationale, and Pustejovsky offers the following explanation:

“Motivations for the postulation of dot objects:

semantic motivations: The knowledge we have of the concepts associated with doors, windows, books, computer programs, etc. is not characterizable as the conjunction

of simple types (or properties) in a conventional type hierarchy. The predicates and relations for the lexical item associated with such a concept are characteristic of that concept alone [...]

lexical motivations: The dot object captures a specific type of logical polysemy, one that is not necessarily associated with true complement coercion [...]”

A dot type is, according to the GL a type of noun that is simultaneously a member of more than one semantic class. According to Rumshisky (2007), the senses—i.e. classes or types—that a dot object presents are metonymically related to one another. This means that the relation between the semantic classes of a dot type is one of regular polysemy. Some examples of dot types are:

k) book : Artifact/Information

l) construction : Process/Result

m) chicken: Animal/Food

n) country: Location/Organization

A dot type selects one or more of its possible senses when placed in a context, as shown in the examples a, b, and c). In case a), *England* selects the Location sense, whereas in case b) it selects the Organization sense. In c) however, the sense of *England* is both “the English organizations” and “the English territory”. We use the name *dot predication* for the instances of a dot type that do not have one of the possible senses as most salient, as in c), which can be seen as a kind of underspecification. Underspecification can be caused by sparseness or by more complex contexts in which both readings are activated at the same time.

In spite of the GL’s computational perspective, Natural Language Processing (NLP) implementations that examine the actual computational feasibility of the GL are few. Moreover, there is no overt attempt to identify the possible three behaviors of a dot type, as the dot predication has not been computationally tackled, which is related to the lack of strategies to capture meaning underspecification.

3. Related work

The work at hand is tangent to the disciplines of Word Sense Disambiguation (WSD), Named Entity Recognition (NER), but also to the NLP-based assessment of lexical semantic theory and to other GL-based NLP tasks.

Joanis and Stevenson (2007) provide a general methodology for NLP tasks aimed towards the validation of lexical semantic theory by means of characterizing the data they want to analyze and correlating the obtained features (in their case, using classification) to the gold-standard provided by the theory. Although their work focuses on verbs and their grain of analysis is at the type level, our work is based on their methodology and adapted to working with nouns at the token level.

A class of nominals that shows regular polysemy and is well studied is the deverbal noun (*destruction*, *examination*), which has distinct grammatical features that can define its reading as either process or result, as covered in theory by Grimshaw (1990) and computationally acknowledged by Peris et al. (2009).

The computational study of regular polysemy has been geared to the collapsing of senses (Vossen et al., 1999; Buitelaar, 1998; Tomuro, 2001) prior to Word Sense Disambiguation (WSD). WSD with collapsed word senses is also known as class-based WSD. The best performance in WSD is obtained by supervised methods that require a very large amount of annotated learning data. Another approach is to use a lexical knowledge base such as a wordnet and a Page-Rank-based algorithm to compute the most likely sense in the sense enumeration of the lexical knowledge base in an unsupervised manner (Agirre and Soroa, 2009). Stanford WordNet (Fellbaum, 1998) does not systematically include metonymical alternations, and the task at hand falls outside the traditional scope of WSD. A lexical knowledge base which does include regular polysemy in its design is the SIMPLE lexicon (Lenci et al., 2000), a GL-compliant lexicon for twelve European languages, as well as other wordnets like DanNet (Pedersen et al., 2009). Class-based WSD however, uses a coarser sense inventory and is more similar to the task at hand, but does not deal with semantic underspecification per se (Izquierdo et al., 2009).

NER, which deals with some cases of regular polysemy, does not deal with semantic underspecification (Sang and De Meulder, 2003). NER shows two different approaches to regular-polysemy based sense alternations, notably for the type Location/Organization. In their account, Johannessen et al. (2005) differentiate what they call the *Form over Function* and the *Function over Form* strategy. Some NER systems assign a constant value to a word type, enforcing what Finkel et al. (2005) call label consistency, namely *Form over Function*. The *Function over Form* strategy, however, assigns a semantic type to the analyzed headword depending on how it behaves in each context and is analogous to the work exposed in this article.

There is also work in the manual and automatic identification of metonymy (Markert and Nissim, 2009) as well as other Generative-Lexicon based sense-disambiguation works, such as Rumshisky et al. (2007) or Pustejovsky et al. (2010). Disambiguation systems, however, are still coping with the need of a representation and recognition of underspecification (Pustejovsky, 2009).

4. Data

Each dataset contains a series of hand-tagged sentences from the ANC. Each sentence has only one token to disambiguate (the *headword*). Each headword can have one of the alternating senses, noted in capitals (LOC vs. ORG, or CTAIN vs. CTENT) or the underspecified sense, which we tagged as DOT.

The lemmas for the Location/Organization dataset were obtained from the ANC from the occurrences of high-frequency (more than 500 occurrences) nouns: Each of the instances was manually identified to obtain their selected sense: Location, Organization or Dot, henceforth LOC, ORG and DOT, following the guidelines detailed in this sections. The list of words for Container/Content was obtained from the dot-type list in Rumshisky (2007).

The data have been annotated by one expert. For any given

instance of a noun X, it was seen if it could be acceptably paraphrased as "the territory of X" (LOC) or "the institutions of X" (ORG). If both applied, it was considered a dot predication (DOT). Likewise, for the Container/Content dot type, the CTAIN and CTENT senses were assigned if the paraphrases "the X as such" or "the content of X" were possible, respectively. If both were possible, the DOT sense was annotated.

After evaluating the SensEval-2007 results, Markert and Nissim (2009) acknowledge the difficulty of identifying specific cases of metonymy for Location and Organization words, and we have considered derivated metonymies from a given class as symptoms of the class itself. For instance, if an Organization type appears very often as a subject, it is very likely to be experiencing the *organization-form-members* metonymy, which we do not separate from the Organization-type behavior, but instead count the presence of the word as agentive subject or perceiver as a potential indicator of its Organization sense.

The *Human Group* reading is metonymically derived from the Organization reading and we treat it as a part of its behavior. This allows us to simplify the Location/Organization/Human-Group postulated by Rumshisky (2007) into a binary dot type Location/Organization. We refer to the first sense listed in the dot type as the *fundamental sense* (LOC, CTAIN) and to the other senses (ORG, CTENT, DOT) as *metonymical senses*.

1. **Location/Organization:** 2132 instances from the words *Afghanistan, Africa, America, Boston, California, Canada, China, England, Europe, Germany, London*.
Sense Distribution: LOC: 1214; ORG: 563; DOT: 355
2. **Container/Content:** 1019 instances from the words *bag, bottle, bowl, box, bucket, container, crate, cup, dish, flask, glass, jar, keg, kettle, pint, plate, pot, spoon, tank, vessel, vial*.
Sense Distribution: CTAIN: 629; CTENT: 144; DOT: 246

The distribution of senses at the word-level is similar to the overall class-wise distribution, that is, words in the Location/Organization dot type tend towards the distribution of senses that the class shows: 57% for LOC, 26% for ORG and 17% for DOT. Some particular words, however, have very different sense distributions. *Afghanistan* selects the Location sense a 75% of the times, whereas *China* selects 48% of its instances as an Organization. We regard these deviations from the overall distribution as the consequence of language use, namely a matter of pragmatics, and not as an intrinsic skewedness in the words' sense selection.

For Container/Content the amount of lemmas is larger than for Location/Organization. During the annotating of the data many examples (around 60%) had to be discarded because they also show irregular polysemy and do not belong to the dot type (e.g. "thick-lensed glasses", "Super Bowl", "army tank").

5. Features

This section describes the feature space used for the characterization of the data listed the previous section. The data has been characterized by means of feature extraction in order to assess the amount of semantic information that their distributional (morphosyntactic and lexical) data can provide.

The features have been extracted from the POS-tagged, XML version of the ANC with noun chunks. No other external resources like FrameNet or WordNet have been used, following Markert and Nissim's (2009) claim that grammatical features—very often subphrasal—tend to be the most discriminating features. For similar remarks, cf. Peris (2009), Rumshisky (2007).

The hypothesis that regular polysemy alternations are often determined at subphrasal level can contradict the idea behind Word Sense Disambiguation (WSD) algorithms like those based on Page Rank, which have a larger, lexical scope of analysis. Selection of metonymically related senses falls outside of the One-sense-per-discourse approach (Gale et al., 1992), since such an approach has been conceived with irregular polysemy in mind, namely for cases like “olive pit” vs. “tar pit”, which depend on contrastive senses.

For each headword token t there is one data example with the following binary features:

1. **NP-traits (6 features)**: these features describe the internal structure of the NP where t appears. The features indicate the presence of an adjective in the NP, of a common noun before or after t , of a genitive mark after t , of a coordinate “X and Y” and the presence of an article in the NP.
2. **Position of t (2 features)**: t being the first or last token of the sentence. This is a coarse approximation of the selected subject position for English (beginning of sentence) or for adjuncts (end of sentence), as no parsing has been used.
3. **Prepositions before t (57 features)**: each feature indicates whether the NP where t is included is introduced by a given preposition. The list of prepositions has been taken from the Preposition Project (Litkowski and Hargraves, 2005).
4. **Previous and next token after t 's NP (4 features)**: each feature describes whether the previous or next token is either a comma or a parenthesis.
5. **Verb after of before t (4 features)**: informs whether there is a verb immediately before t , or whether there is a modal or non-modal verb thereafter.
6. **Lexical space (3000 features)**: A bag of words with 3000 most frequent content words from the ANC.

6. Voting scheme

Our stance is that if the theoretical framework is appropriate for the description of the data, the semantic types it postulates can be validated by being used as target classes in classification experiments. We accept, however, that such

a validation is a costly process because it requires the usage of annotated gold-standard data and is it also subject to the phenomena of sense skewedness, as well as noise and sparseness in the data.

The relevant task of such a system is not to differentiate between the two alternating senses in a dot type (e.g. Location vs. Organization) but rather to automatically identify dot predications (the Dot sense), which fall outside of the design of related Word Sense Disambiguation and Named Entity Recognition work, because a dot predication is an overlapping reading between two orthogonal senses.

The data listed in section 2.1 have been analyzed to obtain the features in 2.2 for each headword, and the resulting set of examples has been used to train and test a classifier. Following Resnik and Bel (2009), we have used Decision Trees because they are more adequate for sparse environments than other families of algorithms (Quinlan, 1993). The classifier was evaluated using a 70-30% split on training and test data.

The baseline approach is to classify the datasets using a Decision Tree, following Resnik and Bel (2010). The results, however, leave room for improvement (cf. Evaluation).

We implemented an ensemble method to improve accuracy. Diversity is a key requisite for any voting scheme (Marsland, 2009), and it was achieved by generating variants of the training sample, as well as using two different learning algorithms.

Variants of the training sample were generated by transforming each tree-class dataset into three binary datasets, in which one selected class is the positive class and the other two become the negative one, tagged together as NO. Thus we obtained the following variant datasets:

1. The original triple set (LOC/ORG/DOT or CTAIN/CTENT/DOT)
2. A set in which the fundamental sense (LOC or CTAIN) is kept, and the rest of senses are tagged as NO.
3. A set in which the metonymical sense (ORG or CTENT) is kept, and the rest of senses are tagged as NO.
4. A set in which the underspecified DOT sense is kept, and the rest of senses are tagged as NO.

We trained a Decision Tree and a KNN (K=21) learner on all variant training sets, which resulted on a total of eight classifiers. We used classifiers from the Orange¹ implementation. All experiments used both datasets, with a 2/3-1/3 split on training and testing. Each example on the test set was classified by each of the eight classifiers, and all the probability distributions were added into one vector with four values. The next example describes the vector of added probabilities for a single test case of the Location/Organization dot type:

$v := [0.88, 1.77, 1.32, 4.02]$

¹<http://orange.biolab.si>

The last value, which is the artificial class NO, was discarded, and then the highest of the other three which are actual classes was used for classification. Here the assigned sense is ORG, which corresponds to the value of 1.77, which is larger than the 0.88 for LOC and the 1.32 for DOT.

7. Evaluation

We evaluated the system on 10-fold cross-validation over a random 2/3 vs. 1/3 data split. The two systems are compared in the next table, which shows the class-wise f-measure for baseline (BL) and voting (VT) systems for Location and Organization (LOC,ORG), Container and Content (CTAIN,CTENT), and their respective dot predications. The voting scheme improved averaged accuracy:

	LOC	ORG	DOT	CTAIN	CTENT	DOT
BL	0.78	0.36	0.48	0.78	0.25	0.32
VT	0.81	0.43	0.31	0.80	0.16	0.32

Table 1: F-measures for each system and class

- a) from 64% to 69% in the Location/Organization dataset over a MFS (Most Frequent Sense) baseline of 56%
- b) from 62% to 66% in the Container/Content dataset over a MFS baseline of 62%

The improvement on performance is significant in both cases (paired t-test gives a p-value <0.001 for both datasets). However, the system does not show overall improvement on the detection of the DOT sense, but rather, it biases the output towards the fundamental sense, which is also the majority class for both datasets, as can be seen from the descent of f-score on classification of metonymical senses, while the f-score for the main sense (LOC or CTAIN) becomes higher with the voting system. This behavior indicates that the system is overfit for the majority class.

The next tables show the compared confusion matrices from the baseline system and the voting system on the right, for the two datasets, in one of the ten experiment runs:

	LOC	ORG	DOT
LOC	248	45	35
ORG	64	73	32
DOT	34	30	43

Table 2: Baseline system on Location/Organization

	LOC	ORG	DOT
LOC	319	29	16
ORG	69	71	29
DOT	39	21	47

Table 3: Voting system on Location/Organization

	CTAIN	CTENT	DOT
CTAIN	145	11	33
CTENT	17	9	17
DOT	35	7	32

Table 4: Baseline system on Container/Content

	CTAIN	CTENT	DOT
CTAIN	177	2	19
CTENT	23	7	13
DOT	53	2	19

Table 5: Voting system on Container/Content

Performance on the Container/Content dataset is lower, because the patterns that pinpoint the senses are less clear-cut than in the LOC/ORG/DOT case. Locations, for example, are very often introduced by prepositions like *in*, and organizations tend to prefer the subject position or the agentive role, which makes them either be the first token of the sentence or the nominal head of a PP introduced by *emphby*. The top features for CTAIN and CTENT are the presence of the preposition *of* after the headword ("a water bottle" vs. "a bottle of water") and the headword being in plural.

8. Conclusions and further work

Dot predications still remain difficult to classify with the current data, features and algorithms. The results suggest that it is required that we measure the ITA (inter-encoder agreement) for this phenomenon, which is very likely to be lower than the expected value of 0.88 provided by Market and Nissim (2007) on their metonymy-resolution work. Measuring agreement also will provide the upper bound for the classification of the datasets. The only available baseline now is MFS baseline.

We also want to measure how the annotation by group of annotators relates to expert-generated data. Ideally, a lower agreement on binary decisions ("Is this a location or an organization?") would hint a DOT sense, as a low ITA can pinpoint semantic underspecification when informants disagree, which can happen in complex contexts, and well as in sparse ones.

On the side of machine learning, resampling can be a good option to relax the MFS baseline, by downsampling the majority class, thereby reducing the statistic bias of the classifier towards the majority class. The following work will compare the results of other classification algorithms.

If we continue to use a voting system, we will have to refine our usage of the artificial class "NO" when establishing the final class of the system. A very sparse data example that consistently gets classified negatively by all the binary classifiers would be a good candidate for a dot predication.

We expect to improve the description of the data by expanding and streamlining the feature system, be it by using other lexical resources or features obtained from other NLP systems like NER. Using parsing instead of only chunking would provide more syntactic information that would complement the morphosyntactic information that the system

already gives account for.

9. Acknowledgements

The research leading to these results has received funding from the European Commission's 7th Framework Program under grant agreement 238405 (CLARA).

10. References

- E. Agirre and A. Soroa. 2009. Personalizing pagerank for word sense disambiguation. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pages 33–41. Association for Computational Linguistics.
- J. D. Apresjan. 1974. Regular polysemy. *Linguistics*.
- S. Banerjee and T. Pedersen. 2002. An adapted lesk algorithm for word sense disambiguation using wordnet. *Lecture notes in computer science*, pages 136–145.
- N. Bel, S. Espeja, and M. Marimon. 2007. Automatic acquisition of grammatical types for nouns. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers on XX*, pages 5–8. Association for Computational Linguistics.
- R. Blutner. 1998. Lexical pragmatics. *Journal of Semantics*, 15(2):115.
- P. Buitelaar. 1998. *CoreLex: systematic polysemy and underspecification*.
- C. Fellbaum. 1998. *WordNet: An electronic lexical database*. MIT press Cambridge, MA.
- Y. Freund and L. Mason. 1999. The alternating decision tree learning algorithm. In *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-*, pages 124–133. Citeseer.
- W. A. Gale, K. W. Church, and D. Yarowsky. 1992. One sense per discourse. In *Proceedings of the workshop on Speech and Natural Language*, page 237. Association for Computational Linguistics.
- J. B. Grimshaw. 1991. *Argument structure*. MIT press Cambridge, MA.
- N. Ide and C. Macleod. 2001. The american national corpus: A standardized resource of american english. In *Proceedings of Corpus Linguistics 2001*, pages 274–280. Citeseer.
- R. Izquierdo, A. Suárez, and G. Rigau. 2009. An empirical study on class-based word sense disambiguation. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pages 389–397. Association for Computational Linguistics.
- Elisabetta Jezek and Valeria Quochi. 2010. Capturing coercions in texts: a first annotation exercise. may. 19-21.
- E. Joanis, S. Stevenson, and D. James. 2006. A general feature space for automatic verb classification. *Natural Language Engineering*, 14(03):337–367.
- J. B. Johannessen, K. Hagen, A. Haaland, A. B. Jónsdóttir, A. Nøklestad, D. Kokkinakis, P. Meurer, E. Bick, and D. Haltrup. 2005. Named entity recognition for the mainland scandinavian languages. *Literary and Linguistic Computing*, 20(1):91.
- A. Kilgarriff, P. Rychly, P. Smrz, and D. Tugwell. 2004. The sketch engine. In *Proceedings of the Eleventh EU-RALEX International Congress*, page 105–116.
- G. Lakoff and M. Johnson. 1980. *Metaphors we live by*. Chicago London.
- M. Lapata and A. Lascarides. 2003. A probabilistic account of logical metonymy. *Computational Linguistics*, 29(2):261–315.
- A. Lenci, N. Bel, F. Busa, N. Calzolari, E. Gola, M. Monachini, A. Ogonowski, I. Peters, W. Peters, and N. Ruimy. 2000. Simple: A general framework for the development of multilingual lexicons. *International Journal of Lexicography*, 13(4):249.
- K. C. Litkowski and O. Hargraves. 2005. The preposition project. In *Proceedings of the Second ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*, page 171–179. Citeseer.
- K. Markert and M. Nissim. 2002. Towards a corpus annotated for metonymies: the case of location names. In *Proc. of LREC*, page 138–1392. Citeseer.
- K. Markert and M. Nissim. 2009. Data and models for metonymy resolution. *Language Resources and Evaluation*, 43(2):123–138.
- B. S. Pedersen, S. Nimb, J. Asmussen, N. H. Sørensen, L. Trap-Jensen, and H. Lorentzen. 2009. Dattet: the challenge of compiling a wordnet for danish by reusing a monolingual dictionary. *Language resources and evaluation*, 43(3):269–299.
- A. Peris, M. Taulp, and H. Rodríguez. 2009. Hacia un sistema de clasificación automática de sustantivos deverbales. *Procesamiento del Lenguaje Natural*, 43:23–31.
- J. Pustejovsky and B. Boguraev. 1996. *Lexical semantics: The problem of polysemy*. Clarendon Press.
- J. Pustejovsky, P. Hanks, and A. Rumshisky. 2004. Automated induction of sense in context. In *Proceedings of the 20th international conference on Computational Linguistics*, page 924. Association for Computational Linguistics.
- J. Pustejovsky, C. Havasi, R. Sauri, P. Hanks, A. Rumshisky, J. Litman, J. Castano, and M. Verhagen. 2006. Towards a generative lexical resource: The brandeis semantic ontology. In *Language Resources and Evaluation Conference, LREC 2006*. Citeseer.
- J. Pustejovsky, A. Rumshisky, J. Moszkowicz, and O. Batiukova. 2009. Glml: Annotating argument selection and coercion. In *IWCS-8: Eighth International Conference on Computational Semantics*.
- J. Pustejovsky, A. Rumshisky, A. Plotnick, E. Jezek, O. Batiukova, and V. Quochi. 2010. Semeval-2010 task 7: argument selection and coercion. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 27–32. Association for Computational Linguistics.
- J. Pustejovsky. 1995. The generative lexicon: a theory of computational lexical semantics.
- J. R. Quinlan. 1993. *C4.5: Programs for Machine Learning*. Morgan Kaufmann.
- G. Resnik and N. Bel. 2009. Automatic detection of non-deverbal event nouns in spanish. *Proceedings of the 5th*

International Conference on Generative Approaches to the Lexicon.

- A. Rumshisky, VA Grinberg, and J. Pustejovsky. 2007. Detecting selectional behavior of complex types in text. In *Fourth International Workshop on Generative Approaches to the Lexicon, Paris, France*. Citeseer.
- N. Tomuro. 2001. Tree-cut and a lexicon based on systematic polysemy. In *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies*, pages 1–8. Association for Computational Linguistics.
- P. Vossen, W. Peters, and J. Gonzalo. 1999. Towards a universal index of meaning. *Proceedings of SIGLEX99: Standardizing Lexical Resources*.
- I. H. Witten and E. Frank. 2005. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann Pub.