

# Identification of Sense Selection in Regular Polysemy Using Shallow Features

**Héctor Martínez Alonso,  
Bolette Sandford Pedersen**  
University of Copenhagen  
Copenhagen, Denmark  
alonso@hum.ku.dk  
bsp@hum.ku.dk

**Núria Bel**  
Universitat Pompeu Fabra  
Barcelona, Spain  
nuria.bel@upf.edu

## Abstract

The following work describes a method to automatically classify the sense selection of the complex type Location/Organization –which depends on regular polysemy– using shallow features, as well as a way to increase the volume of sense-selection gold standards by using monosemous data as filler. The classifier results show that grammatical features are the most relevant cues for the identification of sense selection in this instance of regular polysemy.

## 1 Introduction

In this paper we report on our experiments to automatically assess the distributional evidence that allow the recognition of sense selection for regular polysemy, focusing on the Location/Organization alternation or *dot type* (Pustejovsky, 1995). Broadly speaking, regular polysemy involves the predictable alternation between senses in a systematic way for a significant number of words, i.e. a semantic class or type (cf. section 2). The definition of dot type is further elaborated in 2.1.

The analysis of this data has been implemented by applying a decision tree classifier to the shallow features obtained from a set of occurrences of dot-type words in order to obtain their selected sense. In this aspect, our work is akin to Word Sense Disambiguation (WSD) but it includes an attempt to identify underspecified senses. The machine-learning strategy is also different from state-of-the-art WSD, as seen in sections 3 and 4.

We also propose a method to increase the volume of gold-standard training data by using monosemous words as an aid to provide distribu-

tional information of one of the possible senses in a sense alternation.

The results are expected to give pointers on how to face a general approach for the computational treatment of cases of regular polysemy described as the sense selection of dot types, along with the recognition and tagging of dot predication. The technical application of this research can be used to improve results on information retrieval, semantic role annotation, etc.

## 2 Regular polysemy

*Regular polysemy* is known by several names throughout the literature: *logical*, *complementary* or *systematic polysemy* or even *logical metonymy*.

The wordings are naturally different and may be slightly nuanced, as can be seen by comparing Apresjan's definition (1974, p. 18): "*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."

...with Pustejovsky's definition (1995, p. 28): "*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.*"

From these 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:

- a) Container for content: *He drank a whole glass.*

- b) Property for subject of property: *The authorities arrived quickly.*
- c) Producer for product: *I drive a Honda.*
- d) Location for organization: *France elects a new president.*

This differentiates regular polysemy from what is traditionally referred as polysemy (*irregular polysemy* according to Apresjan), which is more often metaphorical in nature and is sometimes pooled together with homonymy as in the cases of “*olive pit*” vs. “*tar pit*” or “*sand bank*” vs. “*federal bank*”.

### 2.1 Dot type

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 introduces a series of theoretical objects like *qualia structure*, *type coercion* and *dot type*.

The 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:

- e) *book* : Artifact/Information
- f) *construction* : Process/Result
- g) *chicken*: Animal/Food
- h) *country*: Location/Organization

A dot type selects one or more of its possible senses when placed in a context, as shown by the following examples from the American National Corpus or ANC (Ide and Macleod, 2001):

- i) *Manuel died in exile in 1932 in England.*
- j) *England was being kept busy with other concerns*
- k) *England was, after all, an important wine market*

In case i), *England* selects the Location sense, whereas in case j) it selects the Organization sense. In k) 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 k), which can be seen a kind of underspecification.

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 State of the art

The computational study of systematic polysemy has been geared to the collapsing of senses (Vossen et al., 1999; Buitelaar, 1998; Tomuro, 2001) prior to Word Sense Disambiguation (WSD). The best performance in WSD is obtained by supervised methods that require a very large amount of annotated learning data. The other main approach is to use a lexical knowledge base such as WordNet and a Page-Rank algorithm to compute the most likely sense in the sense enumeration of the lexical knowledge base (Agirre and Soroa, 2009). WordNet does not include the Location/Organization alternation in geopolitical locations, so the task at hands falls outside the traditional scope of WSD.

The field of Named Entity Recognition (NER) shows two different approaches to regular-polysemy based sense alternations. 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 word depending on how it behaves in each context and is analogous to the work exposed in this article.

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 help pinpoint its reading as either process or result, as covered in theory by Grimshaw (1990) and computationally acknowledged by Peris et al. (2009).

There is also recent work in the 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).

The SIMPLE lexicon (Lenci et al., 2000) is a GL-compliant lexicon for twelve European languages. It describes its lexical items in terms of their position within a type ontology as well as

a qualia structure. SIMPLE list the *Geopolitical Location* class as a class associated to a complex type  $\langle Location, Human\_Group \rangle$ , which expresses the dot-type ambiguity of words of this class. Words that are considered geopolitical locations can be proper (*Africa, Boston, China*) or common (*city, nation, state*, etc) nouns.

## 4 Experiment

We propose a classification experiment that identifies the senses of Location/Organization words by firstly characterizing the grammatical and lexical features of each and using the extracted features as input for a decision tree classifier. Our experiment can be regarded as a case of WSD in which all disambiguated words can have a Location sense, an Organization sense, or a mixed or underspecified sense which corresponds to the dot predication.

Let  $t$  be the analyzed token of a sentence – the *headword* in WSD jargon–, which belongs to the dot type Location/Organization. The goal of the task is to determine whether each  $t$  has the Location or Organization sense, or rather, if it exhibits a mixed or underspecified behavior, i.e. a dot predication.

The goal of the experiment is to assess the distribution of the complex type Location/Organization and its sense selection in a series of occurrences of proper names for geopolitical locations. A supervised method has been chosen, as this specific phenomenon was expected to require a smaller volume of training data than the general case of supervised WSD.

Markert and Nissim (2009) assume in their metonymy resolution account that the semantic class of the analyzed nouns was already known; claiming that standard NER can be followed by metonymy resolution. We have taken the same assumption and also chosen named entities to build our datasets following their claim that “*Named entities [...] are also very often used figuratively but not listed in dictionaries*”.

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 *org-for-members* metonymy, which we do not separate from the Organization-type behavior, but instead count the presence of

the word as subject as a potential indicator of its ORG sense.

A total of 2132 instances of Location/Organization words were obtained from the ANC from the occurrences of high-frequency (>500) nouns: Each of the instances was manually identified to obtain their selected sense: *Location, Organization* or *Dot*, henceforth LOC, ORG and DOT.

Only one annotator has tagged the data, but Market and Nissim offer a rationale for using one annotator for such coarse-grained distinctions, because they identify an inter-encoder agreement of 0.88. Noise examples (homonymies like *China* being a part of a larger named entity like *AOL China*) were discarded.

For any given instance of a proper noun  $X$ , it was seen if it could be acceptably (albeit possibly in an awkward manner) paraphrased as “*the territory of X*” (LOC) or “*the institutions of X*” (ORG). If both applied, it was considered a dot predication (DOT).

### 4.1 Boosting dataset

The initial distribution of senses is skewed on the side of LOC. In order to balance the Location/Organization distribution of senses, 200 occurrences of *CIA, Microsoft, NATO* and *Pentagon* (also high-frequency words in the ANC) were added because they are purely Organization words that only have Location sense if they experience the *organization-for-headquarters* metonymy, which has not been accounted for. This provides the final distribution of senses in the gold-standard data. It is expected that the Organization sense of the dot types has a similar distributional behavior to the purely Organization-typed word, which allows us to compensate the asymmetry of the data. This has created two different datasets, *Total-dots*, which only has occurrences of words belonging to dot types, and *Total-boost*, which also includes the 800 rows of Organization-type words.

### 4.2 Distribution of senses

The sense distribution is as follows:

	LOC	ORG	DOT	Total
<i>Afghanistan</i>	213	12	60	285
<i>Africa</i>	110	11	18	139
<i>America</i>	69	70	65	204
<i>Boston</i>	121	2	25	148
<i>California</i>	88	16	61	165
<i>Canada</i>	91	43	69	203
<i>China</i>	60	80	27	167
<i>England</i>	86	21	41	148
<i>Europe</i>	151	32	59	242
<i>Germany</i>	123	62	62	247
<i>London</i>	102	6	76	184
<b>Total-dots</b>	<b>1214</b>	<b>355</b>	<b>563</b>	<b>2132</b>
<i>CIA</i>	0	200	0	200
<i>Microsoft</i>	0	200	0	200
<i>NATO</i>	0	200	0	200
<i>Pentagon</i>	0	200	0	200
<b>Total-boost</b>	<b>1214</b>	<b>1155</b>	<b>563</b>	<b>2932</b>

Table 1: distribution of senses for the two datasets

As it can be seen, some lexical elements are much more often Location, which is their fundamental type, but the Organization reading is more common, for instance, for country names. It can also be seen that each lexical item has a different distribution of senses.

## 5 Feature space

The features have been extracted from the POS-tagged, XML version of the ANC with noun chunks, the only source of external information for feature extraction system is the WordSketch (Kilgarrif et al, 2004), which has only been used to establish the nominal word space. No other external resources like FrameNet or WordNet have been used, following Markert and Nissim’s (2009) claim that grammatical features tend to be the most discriminating features. For similar remarks, cf. Peris (2009), Rumshisky (2007).

The hypotheses that regular polysemy alternations are often determined at subphrasal level can contradict traditional WSD algorithms like Page Rank, which have a larger scope of analysis. Selection of metonymical senses falls outside of the One-sense-per-discourse approach (Gale et al., 1992), since such approach has been phrased re-

ferring to irregular polysemy like “*olive pit*” vs. “*tar pit*”.

### 5.1 Lexical and grammatical features

Extracted features are meant to describe the dot type Location/Organization, be it by its grammatical behavior or the lexical environmental that words of this type appear in.

So-called grammatical features describe aspects of the structure of the headword’s NP, its position within the sentence, the relative presence of verbs and punctuations, and most importantly, the presence of prepositions before the headword *t*. Prepositions are regarded as function words and therefore considered part of the grammar.

Lexical features list the words that appear around the instances of the dot type. In order to increase the recall of the system, a set of verbs and nouns from the word sketch of the words *city*, *country* and *continent* –hypernyms for the dot types in the training data– was obtained.

A word sketch is a corpus-based automatic summary of a word’s grammatical and collocational behavior obtained using the Sketch Engine tool (Kilgarrif et al, 2004). Each binary feature informs of the presence of one of the mentioned lemmas in the whole sentence. The verbs were taken from the *object\_of* and *subject\_of* relations, whereas the nouns were taken from the *n\_modifier*, *modifies*, *possession*, *pp\_obj\_of-p* and *pp\_of-p*. Only common nouns have been used. To avoid overfitting the experiment to the sample by using the lexical environment of the analyzed word themselves, the BNC (British National Corpus, distributed by Oxford University Computing Services on behalf of the BNC Consortium) was used. The usage of a lexical environment to assist the disambiguation of a dot type follows Rumshisky (2007).

#### country (noun) bnc freq = 47969

object_of	4871	1.1	subject_of	4322	1.7	adj_subject_of	1046	1.9
flee	<u>70</u>	8.29	ratify	<u>17</u>	6.68	concerned	<u>89</u>	7.34
tour	<u>37</u>	7.51	afford	<u>34</u>	6.62	dependent	<u>21</u>	6.84
visit	<u>102</u>	7.36	border	<u>15</u>	6.54	willing	<u>16</u>	6.36
govern	<u>47</u>	7.22	sign	<u>47</u>	6.29	prepared	<u>18</u>	6.2
rule	<u>47</u>	7.21	experience	<u>34</u>	6.21	capable	<u>8</u>	5.75
leave	<u>332</u>	6.9	benefit	<u>22</u>	6.2	involved	<u>19</u>	5.49
divide	<u>50</u>	6.67	export	<u>15</u>	6.16	rich	<u>10</u>	5.39
represent	<u>94</u>	6.57	suffer	<u>45</u>	6.11	unable	<u>10</u>	5.09
defend	<u>28</u>	6.38	participate	<u>15</u>	6.1	poor	<u>15</u>	4.98

Figure 1: word sketch for “country”

## 5.2 List of features

Following Joanis et al. (2006), the occurrences have been characterized in order to assess the amount of semantic information that their distributional data can provide. The total size of the feature space is of 317 binary features, divided as follows:

1. NP-traits (6 features): which 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 at the beginning of the NP.
2. Position of  $t$  (2 features):  $t$  being the first or last token of the sentence.
3. Prepositions before  $t$  (57 features): each binary feature indicates whether the NP where  $t$  is included is introduced by a preposition. The list of checked prepositions is the one used by the Preposition Project (Litkowski and Hargraves, 2005).
4. Previous and next token after  $t$ 's NP (4 features): each binary feature describes whether the previous or next token is either a comma or a parenthesis.
5. Verb after of before  $t$  (2 features): informs whether there is a verb immediately before  $t$ , or whether there is a modal or non-modal verb thereafter.
6. Lexical space (243 features): The nouns and verbs obtained from the hypernyms’ word sketch.

## 5.3 Classifier runs

In order to establish a classifier, C.45 pruned decision trees from the Weka (Witten and Frank, 2005) implementation were used, as in Resnik and Bel (2009). Decision trees provide an analysis of the importance of the features for a given class, and are more adequate for sparse environments than other families of algorithms (Quinlan, 1993). Due to the relatively small amount of data, performance was evaluated by means of 10-fold cross-validation instead of keeping separate training and test sets.

The six classifier runs can be paired in three groups:

1. *Allthree*: 3-way identification of LOC, ORG and DOT senses from the *Total-dots* and the *Total-boost* datasets.

2. *Loc/Org*: Binary identification of LOC and ORG senses from the *Total-dots* and the *Total-boost* datasets, discarding occurrences tagged as DOT.
3. *Dot/NoDot*: Binary identification of DOT classes from the *Total-dots* and the *Total-boost* datasets, treating both cases of LOC and ORG selection as a NODOT sense.

## 6 Evaluation

The following section details the importance of the lexical features for the construction of the decision tree, as well as the performance measures of the classifier.

### 6.1 Impact of lexical features

This section details the relevance of the lexical features for the decision tree classifier, that is, how relevant the lexical environment is when choosing a possible sense. A very high prevalence of lexical features versus grammatical features would contradict the statement that grammatical features are often key to establish the sense selection of a dot type.

The following tables describe the dimensions of the resulting decision trees for the experiments. *Size of tree* indicates the number of nodes, *number of leaves* is the amount of nodes that assign a sense when reached during the decision process, and *lexical nodes* are those that correspond to one of the 243 lexical features in the feature space. Each binary feature generates two nodes when incorporated into the tree, so 15 lexical items will generate 30 lexical nodes in the decision tree.

	ALL	LOC/ORG	DOT
Size of tree	107	55	65
# leaves	54	28	33
# lexical nodes	38	16	12

Table 2: tree dimensions for Total-boost

	ALL	LOC/ORG	DOT
Size of tree	129	91	51
# leaves	65	46	26
# lexical nodes	30	18	18

Table 3: tree dimensions for Total-dot

The *Total-boost* data, although much larger than the *Total-dots* dataset, generates decision

trees that have a very similar amount of lexical nodes. Lexical nodes are about a third of the total of nodes for a given tree, which always use all the grammatical features such as prepositions and the position of token  $t$  before anything else. Lexical nodes do not appear before the third level in the decision tree, as the first levels are occupied by grammatical features, with the exception of frequent words like *control*, *road* or *south*, which can appear at that level in the different decision trees. Figure 2 shows the first four levels of the LOC/ORG decision trees. The left branch of a given node indicates  $feature=0$  while right branch means  $feature=1$ . The nodes  $p\_in$  and  $p\_from$  and  $p\_to$  represent the features that inform of the presence of the corresponding preposition before the headword.  $L\_paren$  informs of the presence of a left parenthesis,  $NP\_comm$  indicates if there is another common noun in the headword’s NP and *control* is the lexical feature that indicates the presence of the word *control* in the sentence. Underlined nodes are the leaves or output of classifier.

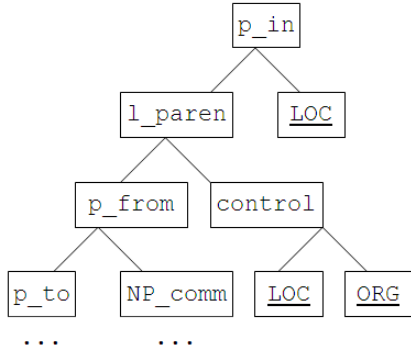


Figure 2: top levels for the LOC/ORG decision tree.

Some prepositions are very safe indicators for LOC, like *in*, *from*, *across*, *over*, while ORG is very often indicated by the relative position of a verb after  $t$  (and therefore a higher likelihood of being a subject) and by prepositions such as *with*, *against*, or *by*, as well as  $t$  being followed by a genitive mark. The contexts that select DOT are much more varied and scarce and the same time, but the prepositions *of* and *for* tend to select for dot predication.

This confirms the position that grammatical elements have more predictive power throughout the datasets than lexical elements for this sort of classification task. The DOT sense is difficult to identify but some prepositions and syntactic contexts favor its unspecified reading. Pure Location/Organization distinction is easy

due to the abundance of fixed syntactic patterns like word order and prepositional cues

## 6.2 Performance

On the account of sparseness, 42 of the lexical features are all-zero, but only 15 rows are all-zero and the rest have at least a feature with a value of 1. Most of the empty lexical features are common collocates for country or city but not for the named-entities which are their hyponyms, like “*country bumpkin*” or “*city dweller*”.

The following tables show the performance of the classifiers in the six runs, the last column lists the Most Frequent Sense (MFS) baseline that the performance is compared against.

	Precision	Recall	F-measure	Accuracy	MFS
Allthree	0,713	0,72	0,72	72%	41%
LOC	0,77	0,8	0,79		
ORG	0,73	0,79	0,76		
DOT	0,55	0,41	0,47		
Loc/Org	0,85	0,85	0,85	85%	51%
LOC	0,86	0,86	0,86		
ORG	0,85	0,85	0,85		
Dot/NoDot	0,57	0,8	0,83	83%	81%
DOT	0,6	0,3	0,4		
NODOT	0,85	0,95	0,9		

Table 4: classifier performance for Total-boost

	Precision	Recall	F-measure	Accuracy	MFS
Allthree	0,69	0,7	0,69	70%	57%
LOC	0,78	0,85	0,81		
ORG	0,57	0,46	0,51		
DOT	0,56	0,53	0,55		
Loc/Org	0,86	0,86	0,86	86%	77%
LOC	0,89	0,94	0,91		
ORG	0,73	0,61	0,67		
Dot/NoDot	0,44	0,76	0,77	77%	74%
DOT	0,6	0,44	0,5		
NODOT	0,82	0,89	0,85		

Table 5: classifier performance for Total-dot

We can see how the *Dot-DoNot* alternation has very poor performance on the DOT class, which is the target class of the experiment. The

DOT class is fuzzier than the other two, and overlaps largely with the ORG class, as it can be seen in the cases of *Allthree* selection. The increase in accuracy in the *Dot-NoDoT* experiment for the *Total-boost* dataset with regards to *Total-dots* is not conclusive, since accuracy is only higher because there are 800 more instances of the most frequent sense, while f-measure remains largely constant for the DOT class in both datasets. The same also holds for the *Allthree* experiment.

In the LOC/ORG case, however, we can see how the *Total-boost* performs substantially better for the ORG class than the *Total-dots*, which is to be expected because it has almost as many examples of each class, while the performance for the LOC class does not become significantly hindered. This supports the usage of boosted datasets as exposed in section 4.1. Accuracy defeats both the expected MFS baseline and the usual baseline for WSD of ~60%.

## 7 Conclusions

It has been seen that the identification of LOC and ORG selection is feasible with good performance measures using only a set of easy-to-obtain linguistic cues and a very naive use of one external resource (WordSketch) which could anyway be replaced by other means of collocation extraction. The experiments confirm the hypothesis that grammatical features are more relevant for the identification of senses in this particular instance of regular polysemy, as lexical items are less represented in the decision trees and seldom appear before the third level of the decision tree. Sparseness means some patterns are underrepresented.

As shown by the performance of the *Total-boost* dataset over *Total-dots*, it is a good idea to increase the volume of sense-selection gold standards by using monosemous data as filler, as this allows training the system on more balanced data. This method can be used to compensate for the skewness of senses in related experiments, as well as to help create faster gold-standard data with a reduced impact on the precision of the system.

The exposed combination of feature space and classifier is suitable for the identification of Location/Organization type selection, but remains insufficient to identify dot predications, although some of them, which are introduced by certain prepositions (*of*, *for*) can be recognized

beforehand and separated from the data in a pre- or postprocessing step.

The poor performance on the identification of dot predication requires a deeper analysis and measure of the inter-encoder agreement for this phenomenon, which is very likely to be lower than the expected value of 0.88 mentioned in section 4.

## 8 Further work

The research has to be expanded to comprise the whole Location/Organization dot type and not only proper nouns, that is, by including common nouns.

After fully studying the Location/Organization dot type, the other listed types (Artifact/Information, etc.) need to be studied in order to grasp the general picture of the soundness of dot type as a theoretical object that can be incorporated into NLP.

More complex, non-shallow features might be necessary for the identification of dot predications. Dependency parsing could indicate some types of dot predications, such as copredications and multiple selections. For other dot types or subtle dot predications, the usage of lexical semantic resources like WordNet might become necessary. The dot type Artifact/Information, for instance, could have a lower inter-encoder agreement than Location/Organization and possibly also a higher relevance of lexical features for the selection of senses, which would imply a smaller role for the grammatical features in the selectional behavior.

An increase in the need to deal with lexical information would also raise the number of features for the sense selection classifiers. Using the ontological types of the words in the lexical feature space instead of the words themselves would reduce the size of the lexical feature space and improve its coverage, as words like *secretary*, *chairman* and *president* would fall under the same ontological type. Ontological types could be obtained from resources like WordNet or the SIMPLE lexicon.

## Acknowledgments

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

## References

- Agirre, E. and Soroa, A. 2009, *Personalizing pagerank for word sense disambiguation*, Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, pp. 33.
- Apresjan, J.D. 1974, *Regular Polysemy*, Linguistics.
- Buitelaar, P. 1998, Phd Thesis, *CoreLex: systematic polysemy and underspecification*.
- Finkel, J.R., Grenager, T. & Manning, C. 2005, "Incorporating non-local information into information extraction systems by gibbs sampling", *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, Association for Computational Linguistics, , pp. 363.
- Gale, W.A., Church, K.W. and Yarowsky, D. 1992, One sense per discourse, Proceedings of the workshop on Speech and Natural Language, Association for Computational Linguistics, pp. 237.
- Grimshaw, J.B. 1991, *Argument structure*, MIT press Cambridge, MA.
- Ide, N. and Macleod, C. 2001, *The american national corpus: A standardized resource of american english*, Proceedings of Corpus Linguistics 2001, pp. 274.
- Joanis, E., Stevenson, S. and James, D. 2006, *A general feature space for automatic verb classification*, Natural Language Engineering, vol. 14, no. 03, pp. 337-367.
- Johannessen, J.B., Hagen, K., Haaland, Å., Jónsdóttir, A.B., Nøklestad, A., Kokkinakis, D., Meurer, P., Bick, E. & Haltrup, D. 2005, "Named entity recognition for the mainland Scandinavian languages", *Literary and Linguistic Computing*, vol. 20, no. 1, pp. 91.
- Kilgarriff, A., Rychly, P., Smrz, P. and Tugwell, D. 2004, *The sketch engine*, Proceedings of the Eleventh EURALEX International Congress, pp. 105–116.
- Lapata, M. and Lascarides, A. 2003, *A probabilistic account of logical metonymy*, Computational Linguistics, vol. 29, no. 2, pp. 261-315.
- Lenci, A., Bel, N., Busa, F., Calzolari, N., Gola, E., Monachini, M., Ogonowski, A., Peters, I., Peters, W. and Ruimy, N. 2000, *SIMPLE: A general framework for the development of multilingual lexicons*, International Journal of Lexicography, vol. 13, no. 4, pp. 249.
- Litkowski, K.C. and Hargraves, O. 2005, *The preposition project*, Proceedings of the Second ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications, pp. 171–179.
- Markert, K. and Nissim, M. 2009, *Data and models for metonymy resolution*, Language Resources and Evaluation, vol. 43, no. 2, pp. 123-138.
- Peris, A., Taulp, M. and Rodríguez, H. 2009, *Hacia un sistema de clasificación automática de sustantivos deverbales*, Procesamiento del Lenguaje Natural, vol. 43, pp. 23-31.
- Pustejovsky, J. 1995, *The generative lexicon: a theory of computational lexical semantics*, MIT press.
- Pustejovsky, J., Rumshisky, A., Plotnick, A., Jezek, E., Batiukova, O. and Quochi, V. 2010, *SemEval-2010 task 7: argument selection and coercion*, Proceedings of the 5th International Workshop on Semantic Evaluation, Association for Computational Linguistics, pp. 27.
- Quinlan, J.R. 1993, *C4.5: Programs for Machine Learning*, Morgan Kaufmann.
- Resnik, G. and Bel, N. 2009, *Automatic Detection of Non-deverbal Event Nouns in Spanish*, Proceedings of the 5th International Conference on Generative Approaches to the Lexicon, .
- Rumshisky, A., Grinberg, V. and Pustejovsky, J. 2007, *Detecting selectional behavior of complex types in text*, Fourth International Workshop on Generative Approaches to the Lexicon, Paris, France.
- Tomuro, N. 2001, *Tree-cut and a lexicon based on systematic polysemy*, Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies, Association for Computational Linguistics.
- Vossen, P., Peters, W. & Gonzalo, J. 1999, *Towards a universal index of meaning*, Proceedings of SIGLEX99: Standardizing Lexical Resources.
- Witten, I.H. and Frank, E. 2005, *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufman.