Combining syntactic patterns and Wikipedia's hierarchy of hyperlinks to extract meronym relations

Debela Tesfaye , Michael Zock and Solomon Teferra
ITPHD PROGRAM, Addis Ababa University, Addis Ababa, Ethiopia
LIF-CNRS, 163 Avenue de Luminy, 13288 Marseille, France
Addis Ababa University, Addis Ababa, Ethiopia
dabookoo@yahoo.com , michael.zock@lif.univ-mrs.fr, solomon_teferra_7@yahoo.com

Abstract

We present here two methods for extraction o, meronymic relation : (a) the first one relies solely on syntactic information. Unlike other approaches based on simple patterns, we determine their optimal combination to extract word pairs linked via a given semantic relation; (b) the second approach consists in combining syntactic patterns with the semantic information extracted from the Wikipedia hyperlink hierarchy (*WHH*) of the constituent words. By comparing our work with SemEval 2007 (Task 4 test set) and WordNet (WN) we found that our system clearly outperforms its competitors.

1 Introduction

The attempt to discover automatically semantic relations (*SR*) between words, or word pairs has attracted a number of researchers during the last decade which is understandable given the number of applications needing this kind of information. Question Answering, Information Retrieval and Text Summarization being examples in case (Turney and Littman, 2005; Girju et al., 2005).

SRs extraction approaches can be categorized on the basis of the kind of information used. For example, one can rely on syntactic patterns or semantic features of the constituent words. One may as well combine these two approaches.

The method using only syntactic information relies on the extraction of word-level, phrase-level, or sentence-level syntactic information. This approach has been introduced by Hearst (1992) who showed that by using a small set of lexico-syntactic patterns (*LSP*) one could extract with high precision hypernym noun pairs. Similar methods have been used since then by (Auger and Barriere, 2008; Marshman and L'Homme, 2006). These authors reported results of high precision for some relations, for example hyponymy, noting poor recall which was low. Furthermore, the performance of this approach varies considerably depending on the type of relation considered (Ravichandran and Hovy, 2002, Girju et al., 2005.

An alternative to the *syntactic approach* is a method relying on the semantics features of a pair of words. Most researchers using this approach (Alicia, 2007; Hendrickx et.al, 2007) rely on information extracted from lexical resources like WN (Fellbaum, 1998). Alas, this method works only for languages having a resource equivalent to *WN*. Yet, even WN may pose a proble because of its low coverage across domains (tennis problem).

Hybrid approaches consist in the combination of syntactic patterns with the semantic features of the constituent words (Claudio, 2007; Girju et.al 2005). They tend to yield better results. However, their reliance on *WN* make them amenable to the same criticism as the ones just mentioned concerning *WN*. More recently Wikipedia based similarity measures have been proposed (Strube, et.al, 2006; Gabrilovich, and Markovitch, 2007). While this strategy produces excellent results, few attempts have been made to extract *SRs* (Nakayama et. al, 2007; Yulan et, al , 2007).

In this paper we propose two approaches to extract meronymic relations. In the first case we rely on the patterns learned from *LSPs*. Previous syntactic approaches aimed at finding stand-alone, unambiguous *LSPs*, for instance *X such as Y*, in order to extract a semantic relation like hyponymy. Yet, such unambiguous, stand-alone *LSPs* are very rare and yield low performance. Instead of using *LSPs* individually, which are often ambiguous, we try to combine them in such a way that they com-

Proceedings of NAACL-HLT 2016, pages 29–36,
San Diego, California, June 12-17, 2016.

plete each other. For instance, the ambiguity of the pattern "NN_1 *make of* NN_2" can be reduced via the pattern "NN_2 *to make* NN_1" in order to extract meronymy. NN_1 and NN_2 can stand for any pair of words. The second approach consists in disambiguating the word pairs extracted by *LSPs* via the information identified from the Wikipedia pages of the respective words.

Our contributions are twofold. First, we propose a novel technique for extracting and combining LSPs in order to extract SRs. Second, we propose an approach for disambiguating the syntactic patterns (say meronymic patterns like NN_1-*has*-NN_2) by building a hyperlink-hierarchy based on Wikipedia pages.

2 Our Approach in more detail

Previous work relies on unambiguous, stand alone *LSPs* to extract SRs. While this approach allows for high precision, it has been criticized for its low accuracy and its variability in terms of the SRs to be extracted. Not all SRs are equally well 'identified'. One of the main challenges and motivations for *LSP* mining lies in the disambiguation of *LSP* to allow for the extraction of SRs. To achieve this, we propose two methods:

- Determine an optimal combination of *LSPs* to represent the relation at hand (section 2.1).
- Combining *LSPs* with the semantic features of the constituent words extracted from the Wikipedia hyperlink-hierarchy (section 2.2).

2.1 Combination of syntactic patterns for relation extraction (CoSP-FRe)

The use of individual *LSP* for the extraction of word pairs linked via a given *SR* tends to produce poor results (Girju et al., 2005; Hearst, 1998). One reason for this lies in the fact that the majority of word pairs are linked via polysemous *LSPs* (Girju et.al , 2005). Hence, these patterns cannot be used alone, as they are ambiguous. At the same time they cannot be ignored as they have the potential to provide good clues concerning certain *SRs*. This being so we suggest to assign weights to the *LSPs* according to their relevance for a specific *SR*, and to optimally combine such weighted patterns for extracting word pairs linked via the SR at hand.

In order to determine the optimal combination of *LSPs* likely to extract *SRs*, we have harvested all *LSPs* encoding the relation at hand. We assigned weights to the patterns according to their relevance for the given *SRs*, and finally filtered the best combination of *LSPs*.

In order to extract such patterns linking word pairs via a certain *SR* , we selected seed-word pairs representative of the relation at hand. In order to balance the word pairs we followed standard taxonomies to group the relations and selected samples from each group (see Section 3.1.1). Sentences containing the word pairs were extracted and then identified their dependency structure. We identified dependency structure linking the word pairs using the shortest path *(ex. nsubj(have, aircraft)* and *dobj(have, engines*) from the *sentence aircrafts have engine)*. Having replaced the words by NN_1 (whole) and NN_2 (part) we obtained patterns like NN_1 *have* NN_2. We finally counted the frequency of the *LSPs* and ordered them according to their frequency and considering the top 50.

Determination of the optimal combination of *LSPs* encoding a given *SR*. To determine the optimal combination of LSPs, we identyfied the discrimination value (*DV*) for each pattern. The *DV* is a numerical value signaling the relevancy of a given *LSP* with respect to a given SR. We applied the following steps in order to identify the *DV* and to determine the optimal combination of the *LSPs*:

Step 1: For each extracted *LSP*, we extracted more connected word pairs from Wikipedia. We defined regular expression matching sentences linking word pairs via the *LSPs* and built then word pairs in a *LSPs* matrix (Matrix 1). Table 1 below shows sample word pairs connected by the patterns NN_1 *has* NN_2 and NN_2 *of* NN_1. Next, we labeled the extracted word pairs with the SR type and built a matrix of word pairs by a specific *SR* type (Matrix 2). In Table 2 the word pairs from matrix 1 are labeled with their respective type of *SR*. We relied on WN to automatically label the word pairs. Starting with the first sense of the words occurring in *WN*, we traverse the hierarchies and identify the *SRs* encoded by the word pairs. Using the information from Matrix 1 and 2, we built a matrix of *SRs* to *LSPs* (Matrix 3). Table 3 shows sample Matrix 3. The rows of the matrix represent the SR type, while columns represent the

LSPs' encoding. The cells are populated by the number of word pairs linked by the *LSP* encoding the *SR*. The *DV* of *LSP* for a given *SR* is given by the following formula:

$$DV = \frac{FPR}{FP} * \log\left(\frac{TNR}{TRE}\right) \quad (1)$$

FP represents the total number of word pairs connected by the *LSP* (from Matrix 1). *FPR* stands for the number of word pairs connected by the given *SR* (from Matrix 2), while *TNR* and *TRE* represent respectively the total number of *SRs* (from Matrix 3) and the total number of SRs encoded by the pattern (from Matrix 3).

Word Pairs	***LSP***
Car Engine	NN_1 has NN_2
Girl Car	NN_1 has NN_2
Door Car	NN_2 of NN_1
Aircraft Engine	NN_1' NN_2

Table 1: Sample Matrix 1.

Word Pairs	***SR* Type**
Car Engine	Meronymy
Girl Car	Possession
Door Car	Meronymy
Aircraft Engine	Meronymy

Table 2: Sample Matrix 2.

***SR* Type**	***NN_1 has NN_2***	***NN_2 of NN_1***	***NN_1' NN_2***
Meronymy	1	1	1
Possession	1	0	0

Table 3: Sample Matrix 3.

Step 2: Identify the optimal combination of *LSP* to represent a given relation. First, we build a matrix combining *LSPs* encoding the respective *SRs* (Matrix 4) from matrix 3. The *LSPs* in Matrix 3 are combined until no other combination is possible. The cells of the Matrix 4 are populated by the number of word pairs linked via the respective combination of *LSPs*. Next we calculated the discrimination value (*DV-g*) for the combined *LSP*s, the *DV-g* being calculated for each combination of *LSP* corresponding to a given SR. We then selected the combination of *LSPs* with maximum *DV-g* for each SR. The *DV-g* for the combined LSPs corresponding to a given SR is given by the following formula:

$$DV - g = \frac{FPR - g}{FP - g} * \log\left(\frac{TNR}{TRE - g}\right) \quad (2)$$

FP-g expresses the total number of word pairs connected by the group of patterns. It is determined by taking the intersection of word pairs connected via the combined *LSPs* (from Matrix 4), where *FPR-g* represents the number of word pairs connected by the combined *LSPs* for a given *SR*. This value is determined by taking the intersection of positive word pairs connected by the combined *LSP* for a given *SR* (from Matrix 4). Finally, *TNR* and *TRE* represent respectively the total number of *SRs* (from Matrix 4) and the total number of SRs encoded by the combination of the *LSP*.

***SR* Type**	***NN_1 has NN_2+ NN_2 of NN_1***	***NN_2 of NN_1+ NN_1' NN_2***
Meronymy	2	2
Possession	0	0

Table 4: Sample Matrix 4.

As can be seen from table 3, the pattern "*NN_1 has NN_2*" when used independently encodes both a meronymic and a non-meronymic word pair. From table 4 above there are two meronymic word pairs linked by the combination of patterns "*NN_1 has NN_2 + NN_2 of NN_1*" while there are no non-meronymic word pairs. Hence the non-meronymic word pair retrieved via the pattern "*NN_1 has NN_2*" is filtered out as a result of having combined it with the pattern "*NN_2 of NN_1*".

2.2 Wikipedia hyperlink hierarchies for *SR* extraction (WHH-Fsre): the case of meronymy extraction

We used here the hyperlink-hierarchies built on the basis of a selected set of sentences of Wikipedia pages containing the respective word pairs in order to disambiguate *LSPs* encoding them. The basic motivations behind this approach are as follows:

1. Words linked to the Wikipedia page title (*WPT*) via *LSP* encoding *SR* are more reliable than word pairs linked in arbitrary sentences.
2. Word pairs encoding a given SR are not always directly connected via *LSPs*. *SRs* encoded by a given word pair can also be encoded by their respective higher/lower order conceptual terms. For instance, the following two sentences "germ is an embryo of seed" and "grain is a seed" yield relations like hyponymy (*germ*, *embryo*, and *grain*, *seed*), meronymy (*embryo*, *seed*, and *germ*, *grain*), the latter (*germ*, *grain*) being inferred via the relation of their higher order terms (*embryo* and *seed*).

The candidate meronymic word pairs extracted via meronymic *LSPs* are further refined by using the patterns learned from their conceptual hierarchies built on the basis of semantic links, namely, '*hypernymic-link'* (*HL*), and the '*meronymic-link'* (*ML*). We extracted the hyperlinks connected to the Wikipedia pages of the respective meronymic candidates by using hypernymic and meronymic *LSP*. The hyperlink hierarchies were built by considering only important sentences (1 and 2 below) from the Wikipedia pages of the pair of terms: (1) definition sentences and (2) sentences linking hyperlinks to the *WPT* using meronymic *LSPs*. Since the meronymic *LSP* vary according to the nature of the arguments, the patterns used to extract hyperlinks for building the hierarchies were learned by taking the nature of the meronymic relations into account (section 2.1). The definition sentences are used to extract *hypernymic-hyperlink*[1], and the sentences linking hyperlinks to the *WPT* using meronymic *LSPs* are used to extract *meronymic-hyperlink*[2]. Using the hierarchy constructed for the candidate word pairs, this approach determines whether the pairs are meronyms or not based on the following assumptions:

(a) The hyperlink hierarchies of hierarchical meronymys constructed form their respective *HL* have a common ancestor in the hierarchy. Figure 1 shows the component-Integral meronyms '*car engine'* sharing the parent '*machine'* in their hyperlink-hierarchy constructed from their respective Wikipedia page definitions.

(b) The hyperlink hierarchies of both hierarchical and non-hierarchical-meronyms constructed from their respective *ML* and/or *HL* converge along the path in the hierarchy.

Extraction of the hyperlinks. To extract the hyperlinks, we performed the following operations:
Step 1: For simple meronymic pairs we identified the respective Wikipedia pages aligning the word pairs with the *WPT* based on the overlap of the surface word form. The word pairs were selected based on standard categories used for describing meronymic taxonomy (Winston et al. 1987, see also section 3.1.1). We first cleaned Wikipedia articles and extracted Wikipedia definitions and sentences linking *WPT* with hyperlinks using meronymic *LSPs*.
Step 2: Annotations. We manually annotated both kinds of sentences using two kinds of information: *WPT* and the hyperlinks. The hyperlink either links the term to its meronyms or hypernyms.
Step 3: Extract *LSPs* linking the *WPT* with the hyperlinks. We assigned *DV* (section 2.1) for the patterns and considered the most frequent *LSPs*. The hyperlinks broadly fall in either of two categories: (a) *hypernymic-hyperlink*. They are extracted by the patterns linking the tuple *(hyperlink, wpt)*, for instance, *is-a (hyperlink, Wikipedia page title)* as in the example (b,c); (b) *meronymic-hyperlinks*. They are extracted via *LSPs* linking the tuple *(hyperlink, wpt)*, for instance, *made-from (hyperlink, wpt)*.

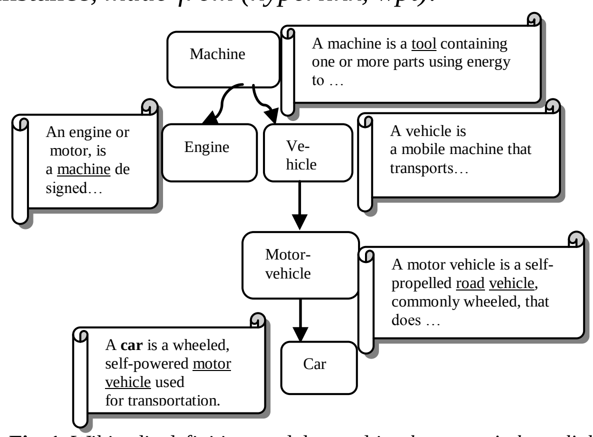

Fig. 1. Wikipedia definitions and the resulting *hypernymic-hyperlink hierarchies* for the meronyms '*car engine'*

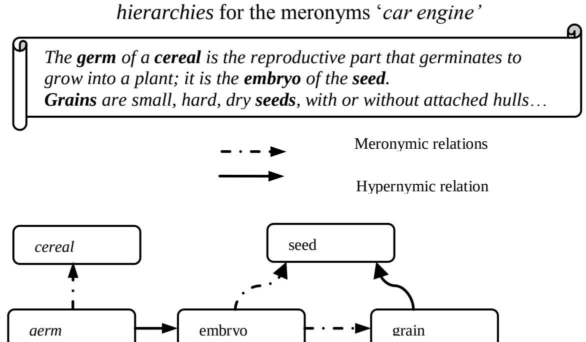

Fig. 2. Wikipedia definitions and the *hyponymic* and *meronymic hyperlink-hierarchies* of the meronym '*grain germ'*

Constructing the hierarchy. For a given pair of terms, we identified the respective Wikipedia

[1] *The hypernymic-hyperlink* is a word defining a term via its higher-order concept, providing in addition a hyperlink to other Wikipedia pages for further reading. The hypernymic-hyperlinks are underlined on figure 1.
[2] *Meronymic-hyperlink* is a word describing a term using its whole concept and providing a hyperlink to other Wikipedia pages for further reading.

pages, by aligning the pairs with the *wpt* and by using word form overlap to extract their associated initial *hypernymic* and *meronymic hyperlinks* (hl_i) based on the patterns learned in step 2.2.1. We further identified the respective Wikipedia pages for the *hypernymic* and *meronymic-hyperlink* (hl_i) identified before and extracted the associated *hypernymic* and *meronymic hyperlinks* (hl_{i+1}). Next we connected (hl_i) with (hl_{i+1}) to form a hierarchy (hypernyms are connected to each other and to meronyms and vice versa). The hyperlinks are extracted until the hierarchies converge, or until the *hypernymic-hierarchy* reaches seven layers (most word pairs converge earlier than that).

Decide on the meronymic status of words. The *hypernymic or meronymic-hyperlink* of one of the words of the pair is searched in the *hierarchy* of the other, and if this link occurs we consider the word pairs as meronyms. Figure 2 shows that the meronymic word pair '*germ grain*' converges at '*seed'* in the *hierarchies* built from their respective Wikipedia pages.

3 Experiment

To show the validity of our line of reasoning we carried out three experiments:

I. Extract the optimal combination of *LSPs* encoding meronymic relation only.
II. Evaluate *CoSP-FRe* for meronymy extraction.
III. Evaluate *WHH-Fre* for extracting meronymy.

3.1 Extract the optimal combination of LSPs encoding meronymy

Training data set. Two sets of data are required: (a) the initial meronymic word pairs used to train our system (b) the corpus from which the *LSPs* were selected. To select the representative list of meronymic pairs, we used a standard taxonomy. Indeed, several scholars have proposed taxonomies of meronyms (Winston et al., 1987; Pribbenow, 1995; Keet & Artale, 2008). We followed Winston's classical proposal:

component – integral-object (cio)	handle– cup
member – collection (mc)	tree – forest
portion – mass (pm)	grain – salt
stuff – object (so)	steel – bike
feature–activity (fa)	paying–shopping
place-area (pa)	oasis–desert

For the training we used the part-whole training set of SemEval-2007 task 4 (Girju et al. 2007) .

Experimental setup. To determine the optimal combination of LSPs encoding meronyms we identified *LSPs* encoding meronymy according to the procedures described in section 2.1. Since most of these patterns are rare we considered only those with a frequency of 100 and above. For individual *LSP* extraction, we identified the *DVs* associated with the meronymic relation by using the formula 1 followed by the *DV-gs* for every combination of *LSPs* by using the second formula. The combined *LSPs* are sorted based on their *DV*. Finally we selected the LSP with the highest *DV* as representatives of the respective meronymic types.

Sno	Pattern	DV-g
1	NN_1 make of NN_2+ NN_2 to make NN_1 + NN_2 used NN_1 + NN_1 NN_2	83.6%
2	NN_1 make from NN_2+ NN_2 to make NN_1 + NN_2 used NN_1 + NN_1 NN_2	81%

Table 5. Part of the optimal combination of patterns for staff object meronymic relations

As can be seen from Table 5 the *DV-g* of staff object meronymic relations patterns is 83.6. The discrimination values for the *LSP* in the group when used individually is below 50%.

Evaluation . The goal is to evaluate the degree of correspondance between the meronyms extracted by *CoSP-FRe* and *WHH-FRe* on one hand and the one by human annotators on the other.

Test data set. We used two data sets: (a) the part-whole test set of the SemEval-2007 task 4 (Girju et al. 2007) which contains 72 examples (26 positive and 46 negative) and some meronymic word pairs gleaned from WN.

Comparison with other systems. We have compared our work against three approaches that achieved the best performance on SemEval-2007 task 4, and two other approaches. We categorized these approaches as (a) *WN*-based: CMU-AT (Alicia, 2007) & ILK (Hendrickx et.al, 2007), (b) syntactic and (c) hybrid approaches: FBK-IRST (Claudio, 2007) & Girjus et.al (2005). We used the individual *LSPs* (*ILSP*) extracted in Sections 2.1 & the *LSPs* extracted by Girju, et.al (2005) as syntactic approach. The *LSPs* extracted by Girju, et.al (2005) are the subset of the LSPs extracted in Sections 2.1.

Results. We computed precision, recall and F-measures as the performance metric. *Precision* is defined as the ratio of the number of correct meronyms extracted and by the total number of extracted word pairs. *Recall* is defined as the ratio between the number of correct meronyms extracted and the total number of meronyms in the test set.

Approaches	**P**	**R**	**F**
CoSP-FRe	76%	88%	81.5%
WHH-FRe	88%	90%	88.9%
ILSP	41.6%	87%	56.2%
CMU-AT	57.7%	45.5%	50.8%
FBK-IRST	65.5%	73.1%	69.1%
ILK	48.4 %	57.7%	52.6%

Table 6. Recall (r), Precision (p) and F-Measure (f) of our approach and related works in the SemEval 2007 test set

We have also extracted meronymic word pairs from random Wikipedia pages of 100 articles and added 85% of the word pairs encoded in *WN*.

Discussions. The results for both approaches are discussed here below:

CoSP-FRe. The precision of *CoSP-FRe* is improved over syntactic approach as the ambiguity of the individual LSP's is reduced when patterns are combined. Recall is improved as a result of using ambiguous *LSPs* for extracting word pairs. This contrasts with all the other syntactic approaches which relied only on unambiguous *LSPs*. In our approach, ambiguous *LSPs* are also used in combination with other *LSPs.* Hence the coverage is significantly improved.

WHH-FRe. Several kinds of hierarchies were formed. Some of them are made of hypernymic or meronymic links, while others are a combination of both links. *WHH-FRe* outperforms significantly previous approaches both with respect to recall and precision as it combines two important features. First *LSPs* are used to extract lists of candidate pairs. Second semantic features of the constituent words extracted from Wikipedia hyperlink-hierarchy is used to further refine. Precision is improved for several reasons: relations encoding *LSPs* which link hyperlinks and *WPT* are more reliable than word pairs connected via arbitrary sentences. The features learned from the Wikipedia hyperlink-hierarchy further cleaned the word pairs extracted by *LSPs*. Recall is also improved since word pairs indirectly linked via their respective higher/lower order hierarchy were also extracted.

4 Related Works

4.1 Syntactic approaches

The work of (Turney, 2005, 2006; Turney and Littman, 2005; Chklovski and Pantel, 2004) is closely related to our work (CoSP-Fre) as it also relies on the use of the distribution of syntactic patterns. However, their goals, algorithms and tasks are different. The work of (Turney, 2005, 2006; and Turney and Littma, 2005) is aimed at measuring relational similarity and is applied to the classification of word pairs (ex. *quart: volume* vs *mile: distance*) while we are aimed at extracting SRs.

4.2 Hybrid approaches

The work of Girju et.al (2005) is more related to our *WHH-FRe* in that they combined *LSPs* with the semantic analysis of the constituent words to disambiguate the *LSPs*. They used *WN* to get the semantics of the constituent words. Alicia (2007) converts word pairs of the positive examples into a semantic graph mapping the pairs to the *WN* hypernym hierarchy. Claudio (2007) combines information from syntactic processing and semantic information of the constituent words from *WN*. Wikipedia-based approaches mainly focused on the identification of similarity (Nakayama et. al, 2007; Yulan et, al , 2007). Also, there is hardly any recent work concerning the extraction of meronyms. Many researchers are working on the identification of semantic similarity achieving excellent result by using standard datasets (Camacho-Collados, Taher and Navigli, 2015; Taher and Navigli , 2015). Yet, most of this work dates back to 2010 and before.

5 Conclusions

We presented here two novel approaches for extracting *SRs*: *CoSP-FRe* and *WHH-FRe*. The strength of *CoSP-FRe* is its capacity to determine an optimal combination of *LSPs* in order to extract *SRs*. The approach yielded high precision and recall compared to other syntactic approaches. *WHH-FRe* perform significantly better than previous approaches both with respect to recall and precision as our approach combines *LSP* and the lexical semantics of the constituted words gleaned from their respective Wikipedia pages.

References

Alain A. and Caroline B. (2008). *Pattern-based approaches to semantic relation extraction: A state-of-the-art.* Terminology Journal, 14(1):1–19

Alicia T. and Scott E. Fahlman (2007). *CMU-AT: Semantic Distance and Background Knowledge for Identifying Semantic Relations*. Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007), pages 121–124, Prague.

Chklovski, T., and Pantel, P. (2004). *VerbOcean: Mining the Web for fine-grained semantic verb relations*. In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-04). pp. 33-40. Barcelona, Spain.

Claudio G., Alberto L., Daniele P. and Lorenza R. (2007). *FBK-IRST: Kernel Methods for Semantic Relation Extraction*. Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007), pages 121–124, Prague.

Fellbaum, C. editor. (1998). *WordNet: An electronic lexical database and some of its applications*. MIT Press.

Gabrilovich, E., Markovitch, S. (2007). *Computing semantic relatedness using wikipedia-based explicit semantic analysis*. In: International Joint Conference on Artificial Intelligence, pp. 12-20.

Girju R., Moldovan D., Tatu, M. & Antohe, D. (2005). *Automatic discovery of Part–Whole relations*. ACM 32(1)

Girju, R., Nakov, P., Nastase, V., Szpakowicz, S., Turney, P., & Yuret, D. (2007). *Semeval- 2007 task 04: Classification of semantic relations between nominals.* In Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval 2007), pp. 13–18, Prague, Czech Republic.

Hearst, M. (1998). WordNet: *An electronic lexical database and some of its applications*. In Fellbaum, C., editor, Automated Discovery of WordNet Relations. MIT Press.

Hearst, M. A. (1992). *Automatic acquisition of hyponyms from large text corpora*. In Proceedings of the 14th International Conference on Computational Linguistics, pages 539–545.

Hendrickx I., Morante R., Sporleder C., Antal v. d. Bosch (2002). ILK: Machine learning of semantic relations with shallow features and almost no data. Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007), pages 121–124, Prague, June 2007

Jose Camacho-Collados, Mohammad Taher Pilehvar ´ and Roberto Navigli (205). *NASARI: a Novel Approach to a Semantically-Aware Representation of Items, Human Language Technologies*. The 2015 Annual Conference of the North American Chapter of the ACL, pp 567–577, Denver, Colorado,USA.

Keet, C.M. and Artale, A. (2008*). Representing and Reasoning over a Taxonomy of Part-Whole Relations.* Applied Ontology, 2008, 3(1-2): 91-110

Marneffe M., MacCartney B. and Christopher D. Manning. (2006). *Generating Typed Dependency Parses from Phrase Structure Parses*. In LREC 2006.

Marshman, E. and M.-C. L'Homme. (2006*).Disambiguation of lexical markers of cause and effect" In Picht, H. (ed.). Modern Approaches to Terminological Theories and Applications*. Proceedings of the 15th European Symposium on Language for Special Purposes, LSP 2005. 261-285. Bern: Peter Lang.

Moldovan D., Badulescu A., Tatu M., Antohe D., and Girju R. (2004). *Models for the semantic classification of noun phrases*. In Proc. of the HLT-NAACL 2004 Workshop on Computational Lexical Semantics, pages 60– 67, Boston, USA.

Nakayama, K., Hara, T., and Nishio S. (2007). *Wikipedia Mining for an Association Web Thesaurus Construction.* In: Web Information Systems Engineering – WISE, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 322-334

Nakayama K., Hara T. and Nishio S. (2008). *Wikipedia Link Structure and Text Mining for Semantic Relation Extraction.* SemSearch 2008, CEUR Workshop Proceedings, ISSN 1613-0073, online at CEUR-WS.org/Vol-334/

Peter D. Turney and Michael L. Littman. (2005). *Corpusbased learning of analogies and semantic rela-tions.* Machine Learning, in press.

Peter D. Turney and Michael L. Littman. (2005). Corpus based learning of analogies and semantic relations. Machine Learning, 60(1–3):251–278

Peter D. Turney. (2006). *Expressing implicit semantic relations without supervision.* In Proceedings of ACL-2006.

Pribbenow, S. (1995). *Parts and Wholes and their Relations*. Habel, C. & Rickheit, G. (eds.): Mental Models in Discourse Processing and Problem Solving. John Benjamins Publishing Company, Amsterdam

Strube, M., and Ponzetto, S.P. (2006). *WikiRelate! Computing semantic relatedness using Wikipedia*. In: Proceedings of the National Conference on Artificial Intelligence, pp. 1419- 1429

Taher M. and Navigli R. (2015). *Align, Disambiguate and Walk: A Unified Approach for Measuring Semantic Similarity.* Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, , Sofia, Bulgaria, 1341–1351, Winston, M., Chaffin, R. & Hermann, D. (1987). *Taxonomy of part-whole relations*. Cognitive Science, 11(4), 417–444.

Yan Y., Matsuo Y. , Ishizuka M. (2009). *An Integrated Approach for Relation Extraction from Wikipedia Texts* . CAW2.0 2009, Madrid, Spain.