Automatic Evaluation of Wordnet Synonyms and Hypernyms

Raghuvar Nadig
Department of Computer Science and Engineering
IIT Bombay
Mumbai, India
raghuvar@cse.iitb.ac.in

J. Ramanand[1]
Cognizant Technology Solutions
Rajiv Gandhi Infotech Park, Hinjewadi
Pune, India
ramanand@it.iitb.ac.in

Pushpak Bhattacharyya
Department of Computer Science and Engineering
IIT Bombay
Mumbai, India
pb@cse.iitb.ac.in

Abstract

In recent times, wordnets have become indispensable resources for Natural Language Processing. However, the creation of wordnets is a time consuming and manpower intensive proposition. This fact has led to attempts at quickly fixing a wordnet using text repositories such as the web and certain corpora, and also by translating an existing wordnet into another language. However, the results of such attempts are often far from ideal, in the sense that the wordnet so produced contains synsets that have outlier words and/or missing words. Additionally, semantic relations may be inappropriately set up or may be missing altogether. This has necessitated investigations into automatic methodologies of wordnet evaluation. This is very much in line with modern NLP's insistence on concrete evaluation methodologies. To the best of our knowledge, the work reported here is the first attempt at an automatic method of wordnet evaluations. We focus on verifying synonymy within non-singleton synsets and also on hypernymy between synsets. Assuming the Princeton WordNet to be the gold standard, our method is shown to validate 70% of all non-singleton synsets and about the same proportion of hypernymy-hyponymy pairs.

1 Introduction

A lexico-semantic network is a graph $<V, E>$, where $v \varepsilon V$ is a concept node and $e \varepsilon E$ is an edge expressing a semantic relation linking two concept nodes. Lexical networks such as the Princeton WordNet (Miller et al., 1990) are now considered vital resources for Natural Language Processing and Text mining. Wordnets are being constructed for different languages: *e.g.,* EuroWordNet (Vossen, 1998), Hindi WordNet (Narayan et al., 2002), BalkaNet (Stamou et al., 2002) *etc*. Competing lexical networks, such as ConceptNet (Liu and Singh, 2004), HowNet (Dong and Dong, 1999), MindNet (Richardson et al., 1998), VerbNet (Kipper-Schuler, 2005), and FrameNet (Baker et al., 1998) have also emerged as alternatives to wordnets.

The construction of lexical networks, however, is a time consuming proposition needing years of effort and substantial linguistic expertise. This fact has prompted attempts at creating them by automatic means using the web or some well structured corpora (*e.g., Wikipedia*[2]*, BNC corpus*[3]*)*, or by translating from an existing network, or through voluntary, participatory effort on the Internet (Senellart and Blondel, 2003). These methods, while fast, often produce resources of dubious quality, in the sense that the concept nodes and semantic relation edges can be inaccurate. In case of wordnets, this lack of quality manifests in the form of outliers or missing words in synsets and/or missing or inaccurate semantic relations.

[1] Work done while at IIT Bombay
[2] http://wikipedia.org
[3] www.natcorp.ox.ac.uk/

Proceedings of ICON-2008: 6th International Conference on Natural Language Processing
Macmillan Publishers, India. Also accessible from http://ltrc.iiit.ac.in/proceedings/ICON-2008

This shows that like other areas of NLP, the question of **evaluation** is important for lexical networks too. One needs criteria and benchmarks for measuring the quality of these structures. As opposed to manual validation, automatic verification techniques to validate lexical networks are advantageous since computers can access a large number of corpus resources and can collate information much faster.

Our literature survey revealed that there have been no comprehensive and principled efforts at evaluating lexical networks, in particular, *wordnets.* The foundational elements of wordnets are **synsets** (sets of synonymous words) and **the hypernymy** (is-a relationship) hierarchy**.** The quality of these two elements ensures the correctness, completeness and the usability of the resource.

Devitt and Vogel (2004) describe a statistical survey of WordNet v1.1.7 to study types of nodes, dimensional distribution, branching factor, depth and height. A syntactic check and usability study of the BalkaNet resource (wordnets in Eastern European languages) has been described in (Smrz 2004). The creators of the common-sense knowledge base ConceptNet carried out an evaluation of their resource based on a statistical survey and human evaluation. Their results are described in (Liu and Singh, 2004). Cuadros and Rigau (2006) discuss evaluations of knowledge resources in the context of a Word Sense Disambiguation task. Cuadros et al. (2007) apply this in a multi-lingual context.

We felt that none of the above addresses the core issues particular to wordnets, and hence we approached the problem by examining *synonymy* and *hypernymy* in wordnets.

This paper is organized as follows: We briefly introduce wordnet synsets and the hypernymy in Section 2. Section 3 deals with the novel problem of synonymy validation and the corresponding experimental results. Section 4 is similarly devoted to hypernymy validation. Finally, Section 5 outlines conclusions and future work.

2 Wordnet synsets and the hypernymy hierarchy

Synsets are the foundations of a wordnet. Obeying the principle of *relational semantics where words disambiguate each other by their mutual association,* a synset is constructed by assembling a set of synonyms that together define a unique sense. The principles of **minimality** and **coverage** (Narayan et al., 2002) are adhered to. For example, in the synset *{stopcock, cock, turncock}*, the word *cock* is highly polysemous, and the unique meaning does not emerge until one of *stopcock* or *turncock* is associated with it. After that, the principle of coverage requires that the remaining word (*stopcock/turncock*) is inserted into the synset. The other determinants of the meaning represented by the synset are (a) the gloss and the example sentence associated with the synset and (b) its semantic linkages with other synsets, especially through hypernymy. The gloss for the synset *{stopcock, cock, turncock}* is *faucet consisting of a rotating device for regulating flow of a liquid*.

A wordnet captures the *"is-a"* relationship - termed *hypernymy/hyponymy* - between synsets. The hypernymy hierarchy of *{stopcock, cock, turncock)* upto a few levels is shown below:

{stopcock, cock, turncock}
=>{hammer, cock}
=> {striker}
=> {mechanical device}
=> {mechanism}
=> {device}

We evaluate the quality of a wordnet by examining the validity of its constituent synonyms and its hypernym-hyponym pairs.

3 Synonymy Validation

We begin the task of synonymy validation by asking the following question: *Are the words in the synset indeed synonyms of each other?* Our algorithm is a method for evaluating this metric. We do not flag omissions in synsets.

Though nothing on synonymy validation, there has been considerable work on *synonymy discovery* (Sanchez and Moreno 2005, Turney 2001, Agirre et al. 2000, Senellart and Blondel 2003). The last mentioned has inspired our work. We illustrate this with an example.

In dictionaries, a word is usually defined in terms of its hypernyms or synonyms. For instance, consider the definitions of the word *snake*, whose hypernym is *reptile*, and whose synonyms are *serpent* and *ophidian* (obtained from the website Dictionary.com[4]):

[4] http://www.dictionary.com

snake: *any of numerous limbless, scaly, elongate* ***reptiles*** *of the suborder Serpentes, comprising venomous and non-venomous species inhabiting tropical and temperate areas.*

serpent: *a* ***snake***

ophidian: *A member of the suborder Ophidia or Serpentes; a* ***snake***.

This critical observation suggests that dictionary definitions may provide useful clues for verifying synonymy.

We use the following observation:

If a word w is present in a synset along with other words $w_1, w_2, \ldots, w_k$, then there is a dictionary definition of w which refers to one or more of $w_1, w_2, \ldots, w_k$ and/or to the words in the hypernymy of the synset.

3.1 Our algorithm for synonymy validation: the basic idea

We apply 3 groups of rules in order. The essential steps in the algorithm are described with examples. For lack of space the fine and exact details have been omitted.

Group 1

Rule 1 - Hypernyms in Definitions

Definitions of words for particular senses often make references to the hypernym of the concept. Finding such a definition means that the word's placement in the synset can be defended.

e.g.

Synset: *{brass, brass instrument}*

Hypernym: *{**wind instrument**, wind}*

Relevant Definitions:

brass instrument: *a musical* ***wind instrument*** *of brass or other metal with a cup-shaped mouthpiece, as the trombone, tuba, French horn, trumpet, or cornet.*

Rule 2 - Synonyms in Definitions

Definitions of words also make references to fellow synonyms, thus helping to validate them.

e.g.

Synset: *{**anchor**, ground tackle}*

Hypernym: *{hook, claw}*

Relevant Definitions:

ground tackle: *equipment, as* ***anchors****, chains, or windlasses, for mooring a vessel away from a pier or other fixed moorings.*

Rule 3 - Reverse Synonym Definitions

Definitions of synonyms may also make references to the word to be validated.

e.g.

Synset: *{Irish Republican Army, IRA, Provisional Irish Republican Army, Provisional IRA, Provos}*

Hypernym: *{terrorist organization, terrorist group, foreign terrorist organization, FTO}*

Relevant Definitions:

Irish Republican Army: *an underground Irish nationalist organization founded to work for Irish independence from Great Britain: declared illegal by the Irish government in 1936, but continues activity aimed at the unification of the Republic of Ireland and Northern Ireland.*

Provos: *member of the Provisional wing of the* ***Irish Republican Army***.

Here *Irish Republican Army* can be validated using the definition of *Provos*.

Rules 4 and 5 - Partial Hypernyms and Synonyms in Definitions

Many words in the wordnet are multiwords, *i.e.*, they are made up of more than one word. In quite a few cases, such multiword hypernyms or synonyms are not entirely present in the definitions of words, but parts of them can be found in the definition.

e.g.

Synset: *{fibrinogen, factor I}*

Hypernym: *{**coagulation factor**, clotting factor}*

Relevant Definitions:

fibrinogen: *a globulin occurring in blood and yielding fibrin in blood* ***coagulation***.

Group 2

Rule 6 – Bag of Words from Definitions

In some cases, definitions of a word do not refer to synonyms or hypernym words. However, the definitions of two synonyms may share common words, relevant to the context of the sense. This rule captures this case.

When a word is validated using Group 1 rules, the words of the validating definition are added to a collection. After applying Group 1 rules to all words in the synset, a bag of these words (from all validating definitions seen so far) is now available. For each remaining synonym yet to be validated, we look for any definition for it which contains one of the words in this bag.

e.g.

Synset: *{**serfdom**, serfhood, vassalage}*
Hypernym: *{**bondage**, slavery, thrall, thralldom, thraldom}*
Relevant Definitions
serfdom: *person (held in)* ***bondage; servitude***
vassalage: *dependence, subjection,* ***servitude***

serfdom is matched on account of its hypernym *bondage* being present in its definition. So the Bag of Words now contains *"person, bondage, servitude"*.
No definition of *vassalage* could be matched with any of the rules from 1 to 5. But Rule 6 matches the word *servitude* and so helps validate the word.

Group 3
Rules 7 and 8 - Partial Matches of Hypernyms and Synonyms
Quite a few words to be validated are multiwords. Many of these do not have definitions present in conventional dictionaries, which make the above rules inapplicable to them. Therefore, we use the observation that, in many cases, these multiwords are variations of their synonyms or hypernyms, *i.e.*, the multiwords share common words with them. Examples of these are synsets such as:

1. *{dinner theater, dinner theatre}*: No definition was available for *dinner theatre*, possibly because of the British spelling.
2. *{Taylor, Zachary Taylor, President Taylor}*: No definition for the last multiword.

Thus the multiword synonyms do share partial words. To validate such multiwords without dictionary entries, we check for the presence of partial words in their synonyms. For example,

Synset: *{Taylor, Zachary Taylor, President Taylor}*
Hypernym: *{President of the United States, United States President, President, Chief Executive}*
Relevant Definitions:
Taylor; Zachary Taylor: *(1784-1850) the 12th President of the United States from 1849-1950.*
President Taylor: - no definition found -

The first two words have definitions which are used to easily validate them. The third word has no definition, and so rules from Group 1 and 2 do not apply to it. Applying the Group 3 rules, we look for the component words in the other two synonyms. Doing this, we find *"Taylor"* in the first synonym, and hence validate the third word.

A similar rule is defined for a multiword hypernym, wherein we look for the component word in the hypernym words. In this case, we would match the word *"President"* in the first hypernym word.

We must note that, in comparison to the other rules, these rules are likely to be susceptible to erroneous decisions, and hence a match using these rules should be treated as a weak match. The reason for the lower confidence in these rules is that these do not validate words directly against evidence external to the wordnet. We can easily create erroneous synsets containing unrelated words which may not have any easily available definitions but share some common words.

The reason for creating these two rules is to overcome the scarcity of definitions for multiwords, and as the coverage of dictionaries improves, the need for these rules will decline.

3.1.1 Experiments on synonymy validation

The validation methodology was tested on the Princeton WordNet (v2.1) noun synsets. Out of the 81426 noun synsets, 39840 are synsets with more than one word, and only these were given as input to the validator. This set comprised of a total of 103620 words.

We created a *super dictionary* which consists of words and their definitions constructed by automatic means from the online dictionary service *Dictionary.com*, which aggregates definitions from various sources. Definitions originating from *Random House* and *American Heritage* dictionaries were included in the dictionary being created. English stop words were removed from the definitions, and the remaining words were stemmed using Porter's stemmer[5]. The resulting dictionary had 463487 definitions in all for a total of 49979 words (48.23% of the total number of words).

3.1.2 Results and discussions on Synonymy validation

Table 1 summarizes the main results obtained by running the dictionary-based validator. This shows that for complete validation, we achieve about 70%

[5] http://www.tartarus.org/martin/PorterStemmer

accuracy. If the criterion is made slightly looser to require that more than half the words are validated, then the figure is impressive, close to 90%.

No. of synsets where all words were validated	27933 (70.11%)
No. of synsets where half or more words were validated	35497 (89.10%)
No. of synsets where none of the words were validated	3660 (09.19%)

Table 1. Statistics on synset validation

In about 9% of all synsets, none of the words in the synset could be verified. Of these 3660 synsets, 2952 (80%) had only 2 words in them. The primary reason for this typically was one member of the synset not being present in the dictionary, thus reducing the number of rules applicable to the other word.

Failure to validate a word does not *necessarily* mean that the word in question is incorrectly present in the synset. Instead, it flags the need for human intervention to verify whether the word indeed has that synset's sense. The next section on case studies shows some interesting examples.

The Princeton wordnet has several words that are not present in conventional dictionaries available on the Web. Encyclopaedic entries such as *Mandara* (a Chadic language spoken in the Mandara mountains in Cameroon), domain-specific words, mainly from agriculture, medicine, and law, such as *ziziphus jujuba* (spiny tree having dark red edible fruits) and *pediculosis capitis* (infestation of the scalp with lice), phrasal words such as *caffiene intoxication* (sic) were among those not found in the collected dictionary. In an evaluation system, the validator can serve as a useful first-cut filter to reduce the number of words to be scrutinized by a human expert.

Rule	*No. of words matching the rule*	*Percentage of total validated words (i.e. out of 85298)*
#1	22845	26.78%
#2	13741	16.11%
#3	7932	09.30%
#4	2378	02.78%
#5	2157	02.53%
#6	905	01.06%
#7	31380	36.79%
#8	3960	04.64%

Table 2. Summary of rule application

3.2 Case Studies

Consider the following cases of words where non-matches are interesting to study. By *non-match,* we mean *flag as not belonging to the synset.*

Instance 1:
Synset: *{visionary,* ***illusionist****, seer}*
Hypernym: *{intellectual, intellect}*
Gloss: *a person with unusual powers of foresight*

The word *"illusionist"* was not matched in this context. This seems to be a highly unusual sense of this word (more commonly seen in the sense of *"conjuror"*). None of the dictionaries consulted provided this meaning for the word.

Instance 2:
Synset: *{bobby pin, hairgrip,* ***grip****}*
Hypernym: *{hairpin}*
Gloss: *a flat wire hairpin whose prongs press tightly together; used to hold bobbed hair in place*

It could not be established from any other lexical resource whether *grip*, though a similar sounding word to *hairgrip*, was a valid synonym for this sense. Again, this could be a usage local to some cultures, but this was not readily supported by other dictionaries.

4 Hypernymy Validation

After synonymy, we undertook the designing of a system for hypernymy validation which takes in two synsets as input and states whether they have a hypernym-hyponym relationship between them. We restrict our work to validating the aforementioned relationship between noun synsets. Like synonymy, our inspiration comes from work on *hypernymy discovery,* though we have not found any work on hypernymy *validation.* Hearst (1992) describes a method of automatic extraction of hy-

pernyms from corpora using syntactic patterns. The work postulates co-occurrence of hypernym-hyponym pairs in syntactic patterns. Hearst's algorithm *Lexico-Syntactic Pattern Extraction* (LSPE) generated six patterns which were used to identify hypernym-hyponym pairs for noun synsets.

Snow et al. (2004) outlined an extension to the above method by training a hypernymy classifier based on dependency trees of known hypernym-hyponym pairs. The classifier was successful in identifying all Hearst patterns, and also gave four new patterns as strong indicators of hypernymy.

Wikipedia[6] has recently inspired extraction of hyponymy relations at a large scale (about 10^6; Sumida and Torisawa 2008). In the following sections, we give our web and Wikipedia based approach to hypernymy validation.

4.1 Our approach: First step: prefix forms as an indicator of hypernymy

Consider the synsets *{racing}* and *{auto racing, car racing}*. Clearly, *auto racing* and *car racing* are specific instances of the term *racing*. A related example is the hypernym-hyponym pair *{work}* and *{paperwork}*.

We generalize these observations to obtain the following rule:
If one term of a synset X is a proper suffix of a term in a synset Y, X is a hypernym of Y

4.2 Second step: using web search to validate hypernymy

For the next step, we use Hearst patterns to validate hypernymy relations between words. The aim is to search for co-occurrences of hypernym-hyponym pairs as Hearst patterns in the corpus. A good corpus for this purpose is the World Wide Web, which offers the following advantages:

1. Its size is larger than that of any other corpus
2. It holds data related to all domains in abundance
3. It has been thoroughly indexed by web crawlers and can be searched easily.

These points above inspire the hypothesis: *If two words in the form of a Hearst pattern show a sufficient number of search results on querying, the words can be validated as coming from a hypernym-hyponym synset pair*.

4.3 Third step: using coordinate terms to validate hypernymy

Consider the following sentence, taken from a hospital brochure:
The first floor houses pediatrics, obstetrics and rheumatology. (*Sentence 1*)

In the above sentence, it is interesting to note that even though the sentence itself contains no reference to a superclass or a hierarchy of terms, it gives us information that there is a degree of similarity in the words *pediatrics, obstetrics* and *rheumatology*, since they are used with the conjunction *and*. An intuitive connection could possibly be that there exists a synset which would be the common hypernym of *pediatrics, obstetrics* and *rheumatology.* According to Princeton WordNet, that synset is *{medicine, medical specialty}*.Thus, *pediatrics, obstetrics* and *rheumatology* are coordinate terms.

Now, if we find Hearst patterns *medicine such as pediatrics* and *medicine such as obstetrics*, we can postulate that *pediatrics* and *obstetrics* are hyponyms of *medicine.* Along with that, we can also go on to postulate that *rheumatology* is also a hyponym of *medicine,* since we have established through Sentence 1 that all three are coordinate terms.

Generalizing, we make the following observation:
If two terms are established to be coordinate terms, a hypernym of one can be stated to be the hypernym of the other.

4.4 Hypernymy validation: Algorithm

The proposed algorithm applies the three steps outlined above to check if the given input synset pairs have hypernymy-hyponymy relationship. First, the synsets are searched for prefix forms. Next, the web search results are utilized for finding co-occurrences of words in both synsets as Hearst patterns (and the patterns obtained in Snow et al, 2004) . Finally, the Wikipedia is used for obtaining coordinate terms of the proposed hyponym terms and the web search step is repeated on them.

Algorithm for validating hypernymy using prefix forms, web search and coordinate terms
1: Input: Synsets S_1 and S_2, web search program, corpus

[6] http://en.wikipedia.org

2: For each combination of words w_1 and w_2 from S_1 and S_2 respectively, check whether w_1 is a prefix of w_2. If yes, label as "validated" and exit.
3: End for
4: Else, For each w_1 and w_2, query the search engine for following terms (with quotes)

i. *"w_1 such as w_2"*
ii. *"such w_1 as w_2"*
iii. *"w_2 and other w_1", "w_2 or other w_1"*
iv. *"w_1 including w2"*
v. *"w_1, especially w_2"*
vi. *"w_1 like w_2"*
vii. *"w_1 called w_2"*
viii. *"w_2 is a w_1", "w_2 is w_1"*
ix. *"w_2, a w_1"*

m = number of above terms with non-zero results
5: If $m >= 2$, label as "validated" and exit
6: End for
7. Else, find coordinate terms w_{2i} for each word w_2 in S_2 and repeat step 4 for each w_1 and w_{2i}

4.5 Experiments

The hypernymy validation procedure was tested on the set of all direct hypernyms for noun synsets in the Princeton WordNet (v2.1)[7]. There are a total of **79297 hypernym-hyponym pairs** constituting this set.

An important aspect of the current work is the assertion of web search results as a tool for the validation of wordnet resources. We run automatic search queries on Microsoft Live search *(http://www.live.com)* and retrieve the number of results obtained. A synset is said to be validated if it gives non-zero search results for any 2 of the 9 patterns tested in the algorithm. This measure is taken to negate the rare occurrences of erroneous Hearst pattern terms, like *canines such as cats*.

The utilization of coordinate terms is achieved by using Wikipedia as corpus. It is searched for syntactic patterns having the forms

1. *w_1, w_2, ...w_{n-1} and w_n*
2. *w_1, w_2, ...w_{n-1} or w_n*

An index is then made using the obtained patterns, such that all words occurring as coordinate terms of a given word can be retrieved. For example, w_1 has w_2, w_3, w_4…..w_n as its coordinate terms because of either of the above patterns.

4.5.1 Results and discussion

The number of pairs validated by the algorithm at different steps of the approach is summarized in Table 3. In all, the algorithm was able to validate 56203 out of 79297 noun hypernymy relation pairs in the Princeton Wordnet, giving a validation percentage of 70.88%. Validation of the hypernymy-hyponymy relation for a pair of synsets is a strong indicator of hypernymy relation between them. However, the failure to validate a synset pair is not a definitive indicator of erroneous construction and has to be treated as a flag for human inspection.

Rule	*Number of hypernym-hyponym pairs validated*	*Percentage of total pairs*
(Rule 1)	16,934	21.35%
(Rule 2)	37,145	46.84%
(Rule 3)	2,124	02.68%
Total	56203	70.88%

Table 3. Results of Hypernymy Validation

5 Conclusion and Future Work

For synonymy validation, our observations show that the intuitive idea behind the algorithm holds well. The algorithm is quite simple to implement and the results easy to interpret. No interpretation of numbers is required; the process is just a simple test. The algorithm is heavily dependent on the depth and quality of dictionaries being used. Since Princeton WordNet is manually crafted by a team of experts, we do not expect to find too many errors. However, many of the words present in the dictionary and not validated were those with rare meanings and usages. Our method makes it easier for human validators to focus on such words. This will especially be useful in validating the output of automatic wordnet creations. The algorithm cannot yet detect omissions from a synset, i.e. the algorithm does not discover potential synonyms and compare them with the existing synset.

The presented algorithm is available only for Princeton WordNet. However, the approach should broadly apply to other language wordnets as well. The limiting factors are the availability of diction-

[7] (http://wordnet.princeton.edu)

aries and tools like stemmers for those languages. Similarly, the algorithm could be used to verify synonym collections such as in Roget's Thesaurus and also other knowledge bases. The algorithm has been executed on noun synsets; they can also be run on synsets from other parts of speech.

For hypernymy validation, our work is based on observations of syntactic patterns shown by hypernyms. The work utilizes Hearst's postulates for formulating a validation approach. Several intuitive ideas are introduced. Web search results are used as a tool for convenient and accurate validation. Coordinate terms are utilized for validation attempts on terms which are rarer in the corpus. Again, the approach is corpus dependent, and a richer corpus would yield better results.

Possible future directions could be expanding the synset validation to other parts of a synset, such as the gloss and relations to other synsets. The results could be summarized into a single number representing the quality of the synsets in the wordnet. The results could then be correlated with human evaluation, finally converging to a score that captures the human view of the wordnet.

We see such evaluation methods becoming increasingly imperative as more and more wordnets are created by automated means.

References

Eneko Agirre, Olatz Ansa, Eduard H. Hovy, and David Martinez. 2000. *Enriching very large ontologies using the WWW.* Proceedings of the Workshop on Ontology Construction of the European Conference of AI (ECAI-00)

Collin F. Baker, Charles J. Fillmore, and John. B. Lowe. 1998. *The Berkeley FrameNet project.* Proceedings of the COLING-ACL

Montse Cuadros and German Rigau. 2006. *Quality Assessment of Large-Scale Knowledge Resources.* Proceedings of Joint SIGDAT Conference on Empirical Methods in Natural Language Processing (EMNLP'06).

Montse Cuadros, German Rigau, and Mauro Castillo. 2007. *Evaluating Large-scale Knowledge Resources across Languages.* Proceedings of the International Conference on Recent Advances on Natural Language Processing (RANLP'07), Borovetz, Bulgaria

Ann Devitt and Carl Vogel. 2004. *"The Topology of WordNet: Some Metrics".* Proceedings of GWC-04, 2nd Global WordNet Conference.

Zhendong Dong and Qiang Dong. 1999. *An Introduction to HowNet.* available from http://www.keenage.com.

Harold P. Edmundson and Martin Epstein. 1969. *Computer Aided Research on Synonymy and Antonymy.* Proceedings of the International Conference on Computational Linguistics.

Marti A. Hearst. 1992. *Automatic acquisition of hyponyms from large text corpora*. Proceedings of the Fourteenth International Conference on Computational Linguistics (COLING), pages 539-545, Nantes, France.

Karin Kipper-Schuler. 2005. *VerbNet: A broad-coverage, comprehensive verb lexicon*. Ph.D. dissertation, University of Pennsylvania

Hguo Liu and Push Singh. 2004. *Commonsense Reasoning in and over Natural Language.* Proceedings of the 8th International Conference on Knowledge-Based Intelligent Information and Engineering Systems.

George A. Miller, Richard Beckwith, Christiane Felbaum, Derek Gross, and Katherine J. Miller. 1990. *Introduction to Wordnet: an on-line lexical database.* The International Journal of Lexicography, vol. 3 (4), pp. 235-244.

Dipak Narayan, Debasri Chakrabarty, Prabhakar Pande, and Pushpak Bhattacharyya. 2002. *An Experience in building the Indo-Wordnet - A Wordnet for Hindi.* Proceedings of the First International Conference on Global Wordnet (GWC '02).

Stephen D. Richardson, William B. Dolan, and Lucy Vanderwende. 1998. *MindNet: acquiring and structuring semantic information from text*. 36th Annual meeting of the Association for Computational Linguistics, vol. 2, pp. 1098-1102

David Sanchez and Antonio Moreno. 2005. *Automatic discovery of synonyms and lexicalizations from the Web.* Proceedings of the 8th Catalan Conference on Artificial Intelligence.

Pierre Senellart and Vincent D. Blondel. 2003. *Automatic discovery of similar words. A Comprehensive Survey of Text Mining,* M. W. Berry, Ed. Springer-Verlag.

Pavel Smrz. 2004. *Quality Control for Wordnet Development.* Proceedings of GWC-04, 2nd Global WordNet Conference

Rion Snow, Daniel Jurafsky, and Andrew Y. Ng. 2004. *Learning syntactic patterns for automatic hypernym discovery* Neural Information Processing Systems, 2004

Sofia Stamou, Kemal Oflazer, Karel Pala, Dimitris Christoudoulakis, Dan Cristea, Dan Tufis, Svetla Koeva, George Totkov, Dominique Dutoit, and Maria Grigoriadou.2002. *BALKANET A Multilingual Semantic Network for the Balkan Languages*. The proceedings of the International Wordnet Conference, Mysore, India, pp. 21-25.

Asuka Sumida and Kentaro Torisawa. 2008. *Hacking Wikipedia for Hyponymy Relation Acquisition.* Proceedings of International Joint Conference on NLP (IJCNLP'08).

Peter D. Turney. 2001. *Mining the Web for synonyms: PMI-IR versus LSA on TOEFL.* Proceedings of the Twelfth European Conference on Machine Learning, Berlin: Springer Verlag, pp. 491-502

Piek Vossen. 1998. *EuroWordNet: a multilingual database with lexical semantic networks.* Kluwer Academic Publishers.