Generating Polarity Lexicons with WordNet propagation in five languages

Isa Maks[1], Ruben Izquierdo[1], Francesca Frontini[2], Rodrigo Agerri[3], Andoni Azpeitia[4], Piek Vossen[1]

[1]VU University Amsterdam [2]Istituto di Linguistica Computazionale, CNR, Pisa,
[3]IXA NLP Group, The University of the Basque Country, San Sebastian [4]Vicomtech-IK4, San Sebastian
{e.maks, r.izquierdobevia, p.vossen}@vu.nl; rodrigo.agerri@ehu.es; aazpeitia@vicomtech.org; francesca.frontini@ilc.cnr.it

Abstract

In this paper we focus on the creation of general-purpose (as opposed to domain-specific) polarity lexicons in five languages: French, Italian, Dutch, English and Spanish using WordNet propagation. WordNet propagation is a commonly used method to generate these lexicons as it gives high coverage of general purpose language and the semantically rich WordNets where concepts are organized in synonym , antonym and hyperonym/hyponym structures seem to be well suited to the identification of positive and negative words. However, WordNets of different languages may vary in many ways such as the way they are compiled, and their numbers of synsets, synonyms and relations. We investigate whether this variability translates into differences of performance when these WordNets are used for polarity propagation. Although many variants of the propagation method are developed for English, little is known about how they perform with WordNets of other languages. We implemented a propagation algorithm and designed a method to obtain seed lists similar with respect to quality and size, for each of the five languages. We evaluated the results against gold standards also developed according to a common method in order to achieve as less variance as possible between the different languages.

Keywords: polarity lexicon, opinion mining, automatic lexicon acquisition

1. Introduction

Sentiment analysis is an active research area including work on acquiring lexica of opinionated words. Most approaches to opinion mining and sentiment analysis rely on these lexicons or lists of words that are used to express sentiment. Knowing the polarity (positive, negative or neutral) of these words helps the system recognize the positive and negative sentiment in the text.

In this paper we focus on the creation of general-purpose (as opposed to domain-specific) polarity lexicons in five languages: French, Italian, Dutch, English and Spanish using WordNet propagation. WordNet propagation is a commonly used method to generate these lexicons as it gives high coverage of general purpose language and the semantically rich WordNets where concepts are organized in synonym , antonym and hyperonym/hyponym structures seem to be well suited to the identification of positive and negative words.

However, WordNets of different languages may vary in many ways such as the way they are compiled, the number of synsets, number of synonyms and number of semantic relations they include. In this study we investigate whether this variability translates into differences of performance when these WordNets are used for polarity propagation. Although many variants of the propagation method are developed for English, little is known about how they perform with WordNets of other languages. We implemented a propagation algorithm and designed a method to obtain seed lists similar with respect to quality and size, for each of the five languages. We evaluated the results against gold standards also developed according to a common method in order to achieve as less variance as possible between the different languages.

The work is carried out within the framework of the FP7 OpeNER project [1]. The main goal of the OpeNER project is to make available a set of open and ready to use tools to perform NLP tasks in multiple languages. As part of this project we developed the polarity lexicons that are described and evaluated in this paper.

The rest of this paper is organized as follows: the first section is a brief overview of related work; section 2 describes the propagation methods and data sets used in this work; in section 3 we present the results of the propagation for the different languages ; section 4 gives an analysis of the results and finally in section 5 some concluding remarks are made.

2. Related Work

Many methods have been developed to generate sentiment lexicons using WordNet as a lexical knowledge base (Liu, 2012). Although these techniques are language and WordNet independent , most have have been tested on the English WordNet only (Esuli and Sebastiani, 2006), (Andreevskaia and Bergler, 2006), (Dragut et al., 2010). These methods differ with respect to how these relations are exploited. Simple techniques start from a few seed words and propagate their polarity through the WordNet. Other methods rely on supervised learning (Esuli and Sebastiani, 2006) or graph-based techniques (Takamura et al., 2006). Our method is close to the simple propagation techniques which start from a seed list and let the list of sentiment words grow by searching the WordNet for synonyms , antonyms and hyponyms. Most of the propagation methods have not been tested on non-English WordNets. We are only aware of a few studies which report results for other WordNets, like (Rao and Ravichandran, 2009) for Hindi, and (Maks and Vossen, 2011) and (Jijkoun and Hoffman, 2009) for Dutch. These studies not only exploit different WordNets but also apply different methods which makes it difficult to compare the results. In the current study we apply one single method

[1]Open Polarity Enhanced Named Entity Recognition (7th Framework Programme, project reference: 296451)

on different WordNets and evaluate the results against comparable gold standards.

3. Data Sets and Propagation Method

3.1. Polarity propagation method

We implemented a propagation algorithm that is similar to propagation methods used by others like (Esuli and Sebastiani, 2006), (Andreevskaia and Bergler, 2006) and (Maks and Vossen, 2011). The method makes use of a seed list of words with known sentiment (positive or negative or neutral). The seed list is semi-automatically aligned with WordNet and the polarity values assigned to the seed synsets are propagated through WordNet using synset relations. This process is carried on (1) until convergence is achieved and no new synsets are added or (2) until max-depth (cf. below) is reached. After assigning the polarity value with highest confidence rate to the synset, a final list of synsets with polarity values is compiled. The following parameters are input parameters of the propagation algorithm:

- max-depth: the maximum level of propagation starting from a seed synset
- relations: a list of synset relations such antonyms, near-synonyms and hyperonyms and an associated weight
- seed list: a file with seed synsets and their polarity values
- WordNet: a file with WordNet data in WordNet-LMF format (Soria et al., 2009)

Optionally, the synset level output is converted to lemma level output where polarity values of the word are computed starting from the polarity of the synsets. In this paper these lemma level versions of the lexicons will be evaluated.

3.2. WordNets

The propagation algorithm is applied on Princeton WordNet for English (Fellbaum, 1998), Cornetto WordNet for Dutch (Vossen et al., 2008), ItalWordNet for Italian (Roventini et al., 2000), WOLF for French (Sagot and Fišer, 2008) and MCR-Spanish WordNet for Spanish (Gonzalez-Agirre et al., 2012).

In this section we present some characteristics which we assume to be relevant for the propagation method. Table 1 presents details of these WordNets : en (English), it (Italian), fr (French), du (Dutch) and sp (Spanish). For each part-of-speech (i.e nouns (n), verbs (v) and adjective (a)) , the table gives the number of synsets and the way the WordNet is compiled, i.e. manually(m), automatically(a) or automatically with manual corrections(a/m). Moreover, it gives the synonym density (syn. dens.) which is the average number of synonyms per synset and the relation density (rel. dens.) which is the average number of 3 types of relations, i.e. near-synonyms, near-antonyms, and hyponyms, per synset. Table 2 gives more details about the distribution of the WordNet relations across the different parts-of-speech.

As can be seen from both tables the WordNets vary with respect to:

		en	fr	it	du	sp
synsets	a	7K	7K	4K	8K	7K
	n	83K	43K	32K	53K	34K
	v	13K	8K	9K	10K	7K
	anv	113K	58K	45K	71K	48K
syn. dens.	anv	1,6	1,7	1,4	1,4	1,5
rel. dens.	anv	1,0	0,9	1,0	1,0	1,0
compiled	anv	m	a	m	m/a	m

Table 1: WordNet characteristcs (K=*1,000)

- size: the number of synsets range from 45,000 (it) to 115,000 (en)
- synset density: the average number of synonyms per synset ranges from 1,4(du, it) to 1,7 (fr).
- compilation: most WordNets are manually built but the Dutch WordNet is partly automatically and partly manually built and the French WordNet is completely automatically built.
- relation density: although the overall average of the number of relations per synset is close to 1,0 for most WordNets (cf. Table 1, row rel. dens.), it differs considerably per part-of-speech across the WordNets. The adjectives seem to have most variation in this respect as the relation density ranges from 0,7 (du) to 2,1 (en).

Below (cf. section 5.) we will see whether these differences across the WordNets have effect on the performance of the method or not.

		en	fr	it	du	sp
rel. dens.	a	2,1	1,8	1,0	0,7	1,7
	n	1,0	0,7	1,0	1,1	0,8
	v	1,1	0,9	1,1	1,1	0,8

Table 2: WordNet density scores per part-of-speech

3.3. Seed lists

The seed lists for the propagation are semi-automatically derived and manually corrected. We use the General Inquirer Lexicon (Stone et al., 1966) as the starting point for the seed lists in all languages. General Inquirer (GI) is a lexicon of English words hand-labeled with categorical information along several dimensions. One such dimension is called valence, with 2,198 unique words labeled 'negative' and 1,915 words labeled 'positive'. We then use the online Google translation service to translate this list of words into Spanish (sp), Italian (it), French (fr) and Dutch (du). For the propagation of the English (en) sentiment lexicons we use the list as it is. In the next step, the synsets of the various WordNets are ordered – from high to low - according to the number of WordNet relations they have. Then, the words of the GI lists are linked to the appropriate synsets and a manual correction is performed for a selection of the first approximately 500 - 700 synsets (i.e. those having most semantic relations) including approx. 150 noun (n), 150 verb (v) and 200 adjective (a) synsets. This process results in a manually corrected seed list of synsets with positive, negative and neutral values (cf. Table 3). Like (Blair-Goldensohn et al., 2008) we start from a seed list that include positive, negative and neutral seeds where the neutral seeds are used to stop the propagation of polarity through synsets with neutral words.

	a	n	v	pos	neg	ntr	all
en	204	180	160	195	222	127	544
fr	187	175	157	192	169	157	519
it	244	227	231	179	194	329	702
du	203	186	174	230	223	109	563
sp	179	186	134	152	102	245	499

Table 3: Seed lists : numbers of synsets

As can be seen from Table 3 the composition of the seed list is quite balanced with respect to the distribution of the seeds across part-of-speech and We chose this method for the composition of the seed list as we showed in a previous study (Maks and Vossen, 2011) that a carefully selected list of seed synsets taking into account the number of lexical relations with other synsets, produces better results than a randomly chosen seed list.

Although the same method was applied to all WordNets, the resulting seed synsets lists have different distributions for positive, negative and neutral synsets. More specifically the Spanish seed list has a low proportion of non-neutral synsets.

Moreover, as can be seen from Table 4, the seed lists differ with respect to both synonym and relation density. For example, the Dutch adjective seed synsets include on average 5,6 synonyms whereas the Spanish ones include only 2 synonyms, on average. Likewise, the Italian seed list has on average 13,5 relation per synset whereas the Spanish seed synsets only have 8,2 relations, on average. Again, the experiments must show whether the results will be effected by these differences. Interestingly, both synonym and relation density of the seed list are much higher than the respective density scores of the whole WordNets (cf. Table 1) where the synonym density score is never higher than 1,7 and the relation density score is never higher than 1,0. The high relation density is due to the method used for the composition of the seed list which aims at a selection of synsets with many relations.The high synonym density score, however, seems to be characteristic of synsets with opinionated words as it supports observations from earlier work (Maks and Vossen, 2010) that opinionated words easily group together in large synsets.

		en	fr	it	du	sp
syn.dens	anv	2.4	2.3	3.6	5.9	2.0
rel. dens.	anv	11,5	9,5	13,5	9,5	8,2

Table 4: Seed lists: density scores

4. Evaluation and Gold Standard

We compiled, for each language, a gold standard in the following way. From the results of a first run of the propagation system, we collected at least 1,000 most frequent words. Frequency is corpus-based using ColFIS for Italian (Bertinetto et al., 2005), D-Coi for Dutch (Oostdijk et al., 2008), BNC for English ((Burnard, 2007), (Gonzalez-Agirre et al., 2012) for Spanish and Wacky for French ((Baroni et al., 2009). We took care to include adjectives(a), nouns(n) and verbs(v) as well as positive(pos), negative(neg) and neutral(ntr) words in order to be able to test reliably positive and negative polarity values across all parts-of-speech. A human annotator was presented with this list and asked to identify the positive, negative and neutral words. The results, for each separate language, are shown in table 5. We think that by composing comparable gold standards, the results of the different WordNets can be compared with each other.

	a	n	v	pos	neg	ntr	all
en	349	493	374	393	342	481	1216
fr	321	458	338	320	324	473	1117
it	400	304	300	311	172	521	1004
du	387	381	269	351	267	419	1037
sp	465	638	614	462	381	869	1717

Table 5: Composition of the gold standards

5. Experiments and Results

We evaluated several versions of our method in order to study the behaviour of the different WordNets in combination with the different settings. The results are evaluated against the gold standards described in section 4. We evaluated the resulting selections of positive and negative words in terms of precision. We excluded neutral words from the lexicon selections as we are not interested in their performance. However, the gold standard includes neutral words to be able to find words false positive cases, i.e. words which are incorrectly labeled as neutral instead of positive or negative. As the gold standard is a subset of the generated lexicons we do not present recall scores. Instead, we give the number of lemmas (cf. lsize).

5.1. Zero iterations

In the first setting we run the method with zero iterations generating lexicons which include only those words which are member of the seed synsets. In this case, the method generates lemma level output from synset level output without any form of propagation. As can be seen from Table 6 precision ranges from 79% to 89% which means that going from synset to lemma causes a considerable loss of performance (cf. section 5.2.2. below0. Interestingly, although the seed lists of the different languages are of comparable size, the size of the zero iterations lexicons varies from 467 (sp) to 2,740 (nl) lemmas. This is the direct effect of the differences between the seed lists with respect to synonym density already noted in the previous section 3.3.: the Dutch and Italian seed list have a high synonym density and therefore generate large lexicons whereas the Spanish seed list has a relatively low synonym density and thus generates a much smaller lexicon. We consider these lexicons as baselines as - with our method - precision can never be higher than the baseline precision and the lexicon size can never be smaller.

5.2. Different types of WordNet relations

Table 6(row 2-5) presents the results of the propagation with one type of semantic link in isolation: near-synonym(syn), near-antonym (ant),hyponym (hypo) and hypernym (hype) links. The next rows (5-8) give the results of the propagation with combinations of links: near-synonym, near-antonym and hyponym (sahypo), near-synonym, near-antonym and hypernym (sahype), near-synonym, near-antonym, hyponym and hypernym(sahohe). Near-synonym and near-antonym relations generate lexicons with high precision. However, as there only few of them (cf. Table 2) they generate small lexicons. Interestingly, although the selections generated with hyperonym relations are considerably smaller than the selections with hypernym relations (cf. row sahype vs. sahypo), they often have similar precision. This implies that hyponymy is a better link for propagating polarity through the WordNets than hyperonymy. The combination of all relations (saheho) generates the largest lexicons with, however, low precision in most cases.
Although the results of the various WordNet differ considerably with respect to precision we do not see a clear relation between the synset characteristics of the WordNets the reported results. The Spanish lexicon is relatively small and this may be due to the low relation density of the seed list. However, the English WordNet has similar scores for relation density and yet generates the largest lexicons. We consider the combination of near-synonym, near-antonym, and hyponym (sahypo) relations as best as it generates sufficiently large lexicon with relatively high precision. We use these settings for further experiments reported in the following sections.

5.2.1. Confidence Levels

Our method generates a confidence score for each synset with a polarity value. The score is calculated as a combined measure of (1) the length of the path from the synset to the the seed synset and (2) the weight of the wordNet relations that are part of the path. WordNet relations have a weight that can be set before running the program. In the current experiments, the weights are set as follows: near-synonyms and near-antonyms relations both have a weight '2' and hyponym relations have weight '1'. Table 7 presents results of the selections which take into account the various confidence levels. With the heightening of the level (cf. column 1 'conf'), the size of the lexicon decreases and precision increases. Interestingly, precision on these selections is higher than precision on selections of comparable size achieved by the propagation with limited sets of links (cf. 6, row 2-5). For example, a selection of the English results with a confidence level higher than 0.2 consists of 5,705 words with an precision of 77% whereas a selection of similar size achieved by propagation with near-synonym, antonym and hyperonym links (cf. 7, column 1 , row 7) has considerably lower precision (72%). Likewise, for Spanish , results with a confidence level higher than 0.3 (cf.7,) outperform with 90% all other results achieved for Spanish. We conclude therefore that it is better to generate small but high quality lexicons with the heightening of the confidence score than, for example, with the choice of high-scoring relations.

5.2.2. Synset to word approach

The final step involves the generation of a lemma level lexicon where each word has one polarity value. The advantage of such a lexicon is that it can be used with opinion mining tools that do not rely on deep semantic analysis or word sense disambiguation. The computation of the polarity of a lemma from the polarity of the synset can be done in various ways (cf. (Gatti and Guerini, 2012)). First, all the polarity values of each sense together with their confidence rate are collected; then we obtain the overall lemma's polarity by following one of the following 3 heuristics:

- majority: the word gets the polarity value which is assigned to the majority of its senses
- average: the word gets the polarity value with on average has the highest confidence rate
- maximum: the word gets the polarity value with the highest confidence rate

Tables 8 and 9 present the results of the different settings for generating word lexicons from the synset results. For all languages and both experiments , the 'majority' approach outperforms the other ones. We used the 'majority' approach in all other experiments reported in this paper.

	avg	maj	max
en	.68	.69	.66
fr	.59	.62	.59
it	.63	.72	.63
du	.65	.69	.66
sp	.67	.79	.67

Table 8: Results of synset to word heuristics (sahypo)

	avg	maj	max
en	.89	.89	.85
fr	.81	.80	.80
it	.80	.80	.73
du	.81	.80	.80
sp	.80	.79	.76

Table 9: Results of synset to word heuristics (baseline)

The baseline scores (cf. Table 9) probably show best the 'damaging' effects of this synset to word step. As the baseline score is based on the evaluation of the *synset* level seed list against the *word* level gold standard, it basically

	en		fr		it		du		sp	
	pr	lsize	pr	lsize	pr	lsize	pr	lsize	pr	lsize
baseline	.89	890	.80	681	.80	1323	.81	2740	.81	467
syn	.79	3695	.80	1309	.80	1590	.79	3151	.81	807
ant	.88	1143	.77	841	.78	1630	.78	3094	.81	554
hypo	.71	11,475	.57	4,338	.76	4891	.72	10,851	.70	3097
hyper	.76	1282	.74	868	.76	1395	.71	3231	.72	569
sahypo	.69	16,569	.62	5416	.72	5465	.69	12,553	.79	3626
sahyper	.72	5296	.76	1760	.72	2071	.69	4060	.79	1045
saheho	.64	40167	.57	9828	.65	11,460	.58	22,846	.76	5649

Table 6: Results with different WordNet relations (ls=lexicon size ; pr = precision)

	en		fr		it		du		sp	
	pr	lsize	pr	lsize	pr	lsize	pr	lsize	pr	lsize
0.0	.69	16,569	.62	5416	.72	5456	.69	12,553	.81	3626
0.1	.70	13,596	.65	4595	.72	5014	.69	11,577	.80	3484
0.2	.77	5705	.78	1605	.74	2188	.76	4912	.84	2048
0.3	.78	4803	.81	1257	.77	1915	.76	4062	.90	1814

Table 7: Results with different confidence levels (ls=lexicon size ; pr = precision)

measures the performance of the synset-to-word step which comes with a loss of 11% to 20% . However, in the next session can be seen that this drop in performance does not occur with all parts-of-speech.

5.3. Part-of-speech

We saw that the distribution of the synonyms across synsets (i.e. synonym density) and the distribution of WordNet relations across synsets (i.e. relation density) differ by part-of-speech. In this section we analyze the results for each separate part-of-speech to discover whether there is a correlation between the WordNet's density scores and the propagation results. Table 10 presents precision scores per parts-of-speech, i.e. adjectives(a), nouns(n) and verbs(v).

In almost all cases, precision on adjectives is considerably higher than precision on the other parts-of-speech. We assume that polysemy has more impact on verbs and nouns than on adjectives. Although adjectives can be highly polysemous , usually most senses have the same polarity. We can see this trend with both baseline scores (ranging from 88% to 99%) and the 'sahypo' selection.

The results for the other parts-of-speech differ considerably across the different WordNets but are generally low, except for the English nouns.

		en	fr	it	du	sp
a	baseline	.99	.88	.88	.87	.93
	sahypo	.76	.81	.75	.77	.93
n	baseline	.90	.76	.67	.77	.25
	sahypo	.77	.64	.65	.57	.38
v	baseline	.77	.71	.62	.72	.72
	sahypo	.54	.38	.48	.65	.67

Table 10: Results per part-of-speech

6. Analysis and Discussion

Our main question was whether the differences between the WordNets would effect the performance of the propagation method. We therefore reported statistics and other characteristics of both the WordNets (cf 3.2.) and the seed lists (cf.3.3.) and tried to relate these with the results of the propagation method. We repeat the most important characteristics of the WordNets (cf. Table 11) which are the way in which the WordNet is compiled (first row), the synonym density , the relation density and the size of the WordNet in number of synsets . Additionally we give the sum of the relation density and the synonym density of the seed list (seed list density) . The lower part of the Table shows the size, in number of synsets, and the precision of the 'best' lexicon obtained by propagation (sahypo, cf. 5.2.) .

	en	fr	it	nl	sp
compiled	m	a	m	m/a	m
synonym density	1,0	0,9	1,0	1,0	1,0
relation density	1,6	1,7	1,4	1,4	1,5
number of synsets	113k	58k	45k	71k	48k
seed list density	13,9	11,6	16,1	15,4	10,2
size (sahypo)	16k	4k	4k	12k	4k
precision (sahypo)	.69	.62	.72	.69	.79

Table 11: WordNet characteristics vs. propagation results (k=*1000)

We observe that the resulting lexicons vary considerably in size, ranging from 4K to 16K and that the largest WordNets (cf. 'en' and 'nl') generate the largest lexicons. Moreover, we observe that precision scores of the generated lexicons vary greatly. We expected to find a correlation between the the WordNet density and the precision scores, but this seems not to be the case. The lexicon with the lowest score ('fr') has the highest relation density score and WordNets with equal relation density score generate lexicons with different precision scores ('du' with 0.69 vs. 'it' with 0.72). Moreover, not even a high seed list density score contributes to the precision of the generated lexicon as the best scoring lexicon (sp) is generated by a seed list with the lowest density and moderate scoring lexicons ('en' and 'nl') are generated by seed lists with relatively high density. However, there seems to be a relation between the way the WordNets are built and the precision scores: probably the low scores of the French WordNet must be explained by the fact that it is completely automatically compiled. To conclude, although the propagation method generates results with different precision scores for each WordNet, we cannot find other WordNet characteristics than the way the WordNet is compiled to explain these variations.

However, there are some other interesting observations to make which relate to the similarities - rather than to the dissimilarities - across the results.

First, although we used the best performing heuristic for generating a lemma level lexicon from a synset level lexicon , this step causes a considerable drop in performance in all cases suggesting that word polarity is quite different from synset polarity.

We also observed that in almost all cases larger selections of words have considerably lower performance than smaller selections and that propagation comes with loss of accuracy. The best selections are achieved by applying several links together and then filtering the results by using the confidence level. This shows that taking into account all information available helps to improve the propagation results more than just focussing on high-scoring WordNet relations.

Finally, with almost all WordNets nouns and verbs perform rather poorly and are greatly outperformed by adjectives. This suggests that the organization of adjectives in synsets which are related to each other by near-synonym and antonym relations, is more effective when looking for polarity than the organization of nouns and verbs in hyponym/hypernym structures. We also assume that, in the case of polarity identification, polysemy has higher impact on verbs and nouns than on adjectives.

7. Conclusion

In this paper we investigated whether WordNets of different languages give similar results when used to generate polarity lexicons by WordNet propagation. We found that, although the quality of the results differs greatly, this could not be related to any of the relevant WordNet characteristics except for the way it is compiled. More specifically, a fully automatically compiled WordNet seems not to be appropriate for polarity propagation. Moreover, we observed some interesting phenomena occurring across all WordNets used in this study. They suggest that the propagation method should be primarily used to identify polarity of adjectives rather than polarity of verbs and nouns.

8. Acknowledgements

This work is part of the OpeNER project funded by the European Commission 7th Framework Programme (FP7), grant agreement no 296451.

9. References

Andreevskaia, A. and Bergler, S. (2006). Sentiment tagging of adjectives at the meaning level. In *Canadian Conference on AI*.

Baroni, M., Bernardini, S., Ferraresi, A., and Zanchetta, E. (2009). The wacky wide web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation*, 43(3):209–226.

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., and Thornton, A. M. (2005). Corpus e lessico di frequenza dell'italiano scritto (colfis).

Blair-Goldensohn, Sasha, Hannan, Kerry, McDonald, Ryan, Neylon, Tyler, Reis, George A., and Reynar, Jeff. (2008). Building a sentiment summarizer for local service reviews. In *Proceedings of WWW-2008 workshop on NLP in the Information Explosion Era*.

Burnard, Lou. (2007). Reference guide for the british national corpus (xml edition). Technical report, Research Technologies Service at Oxford University Computing Services.

Dragut, E., Yu, C., Sistla, P., and Meng, W. (2010). Construction of a sentimental word dictionary. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*.

Esuli, Andrea and Sebastiani, Fabrizio. (2006). Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of LREC2006*, Genova, Italy.

Fellbaum, Chr. (1998). *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA, USA.

Gatti, Lorenzo and Guerini, Marco. (2012). Assessing sentiment strength in words prior polarities. In Kay, Martin and Boitet, Christian, editors, *COLING (Posters)*, pages 361–370. Indian Institute of Technology Bombay.

Gonzalez-Agirre, A., Laparra, E., and Rigau, G. (2012). Multilingual central repository version 3.0: upgrading a very large lexical knowledge base. In *Proceedings of the 6th Global WordNet Conference (GWC'12)*, Matsue, Japan.

Jijkoun, V. and Hoffman, K. (2009). Generating a non-english subjectivity lexicon: relations that matter. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistic 2009 (EACL)*, Athens, Greece.

Liu, Bing. (2012). *Sentiment Analysis and Opinion Mining*. Morgan and Claypool Publishers.

Maks, Isa and Vossen, Piek. (2010). Modeling attitude, polarity and subjectivity in wordnet. In *Proceedings of the 5th Global WordNet Conference (GWC'10)*, Mumbai, India.

Maks, Isa and Vossen, Piek. (2011). Different approaches to automatic polarity annotation at synset level. In *Proceedings of the 1st International ESSLLI Workshop on Lexical Resources (WoLeR 2011)*, Ljubljana, Slovenia.

Oostdijk, N., Reynaert, M., Monachesi, P., van Noord, G., Ordelman, R., Schuurman, I., and Vandeghinste, V. (2008). From d-coi to sonar: A reference corpus for dutch. In *Proceedings of LREC2008*, Marrakech, Morocco.

Rao, D. and Ravichandran, D. (2009). Semi-supervised polarity lexicon induction. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistic 2009 (EACL)*, Athens, Greece.

Roventini, A., Alonge, A., Calzolari, N., Magnini, B., and Bertagna, F. (2000). Italwordnet: a large semantic database for italian. In *Proceedings of LREC2000*, Athens, Greece.

Sagot, B. and Fišer, D. (2008). Building a free french wordnet from multilingual resources. In *Proceedings of Ontolex 2008*, Marrakech, Morocco.

Soria, C., Monachini, M., and Vossen, P. (2009). Wordnet-lmf: fleshing out a standardized format for wordnet interoperability proceedings of iwic2009, stanford, usa, february 20-21, 2009, p. 139-146, ed. acm press, new yo. In *Proceedings of IWIC2009*, Stanford, USA.

Stone, Philip, Dunphy, Dexter, Smith, Marshall, and Ogilvie, Daniel. (1966). *The General Inquirer: A computer approach to Content Analysis*. MIT Press.

Takamura, Hiroya, Inui, Takasha, and Okumura, Monabu. (2006). Latent variable models for semantic orientation of phrases. In *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistic 2006 (EACL)*.

Vossen, P., Maks, I., Segers, R., and van der Vliet, H. (2008). Integrating lexical units, synsets and ontology in the cornetto database. In *Proceedings of LREC2008*, Marrakech, Morocco.