Extending wordnets by learning from multiple resources

Benoît Sagot, Darja Fišer,

Extending wordnets by learning from multiple resources

Benoît Sagot[+], Darja Fišer[*]

[+]Alpage, INRIA & Université Paris 7, 175 rue du Chevaleret, 75013 Paris, France
[*]University of Ljubljana, Department of Translation, Aškerčeva 2, SI – 1000 Ljubljana
benoit.sagot@inria.fr, darja.fiser@ff.uni-lj.si

Abstract

In this paper we present an automatic, language-independent approach to extend an existing wordnet by recycling existing freely available bilingual resources, such as machine-readable dictionaries and on-line encyclopaedias. The approach is applied to Slovene and French. The words extracted from the bilingual resources are assigned one or several synset ids based on a classifier that relies on several features, including distributional similarity. Automatic and manual evaluation shows that the resulting extensions of sloWNet and WOLF are lexico-semantic repositories of high coverage as well as high quality.

Keywords: WordNet extension, WOLF, sloWNet, word sense disambiguation, distributional semantic models

1. Introduction

As the role of lexical knowledge is gaining importance in many areas of natural language processing, several frameworks for organizing and representing it have been proposed, such as ACQUILEX, MindNet, ConceptNet or Cyc. One of the best-known and most widely used lexico-semantic resources is Princeton WordNet (Fellbaum, 1998) and its sister wordnets for languages other than English, such as EuroWordnet (Vossen, 1999) and BalkaNet (Tufiş, 2000).

Because manual construction of such resources is too time-consuming and expensive to be feasible for most research scenarios, semi- or fully automatic approaches have recently become popular, which exploit various types of existing resources to facilitate the development of a new wordnet. However, a common problem with automatically induced wordnets is the necessary trade-off between a limited coverage and the desired level of accuracy, both of which are required if the resource is to be useful in a practical application.

We present here an approach for extending existing wordnets by extracting additional lexico-semantic information from already available bilingual language resources and then training a maximum entropy classifier on the existing core wordnet in order to assign the new vocabulary to the appropriate synsets. Our approach, applied on the French wordnet WOLF and the Slovene wordnet sloWNet, handles monosemous and polysemous words from all parts of speech.

This paper is structured as follows: in Section 2 we give an overview of related work. In Section 3 we introduce WOLF and sloWNet. In Section 4, we describe the process of extracting lexico-semantic information from bilingual lexical resources. In Section 5 we explain the wordnet enrichment experiment using a maximum entropy classifier that helped us determine whether a translation we extracted from the existing resources is an appropriate candidate for a given synset. Section 6 is dedicated to the evaluation of the extended resources.

2. Related work

Most automatic approaches to create a wordnet for a new language take the Princeton wordnet as a backbone and extend it with the vocabulary inventory of the target language. Among the most straightforward resources to obtain lexical knowledge for the language in question are machine-readable bilingual dictionaries. Entries from the dictionary are linked to PWN synsets under the assumption that their counterparts in the target language correspond to the same synset (Knight and Luk, 1994). Obviously, bilingual dictionaries are generally not concept-based but follow traditional lexicographic principles, which is why the biggest obstacle is the disambiguation of dictionary entries.

This problem is overcome by a different set of approaches in which bi- or multilingual lexicons are extracted from parallel corpora (Fung, 1995). The main underlying assumption in these approaches is that senses of ambiguous words in one language are often translated into distinct words in another language (Dyvik, 2004; Ide *et al.*, (2002). Furthermore, if two or more words are translated into the same word in another language, then they often share some element of meaning. This results in sense distinctions of a polysemous source word or yields synonym sets.

The third set of approaches that have become popular in the past few years draw upon Wikipedia. New wordnets have been induced by using structural information to assign Wikipedia categories to WordNet (Ponzetto and Navigli, 2009) or by extracting keywords from Wikipedia articles (Reiter *et al.*, 2008). Vector-space models to map Wikipedia pages to Wordnet have been developed (e.g., by Ruiz-Casado *et al.*, 2005). The most advanced approaches use Wikipedia and related projects, such as Wiktionary, to bootstrap wordnets for multiple languages (Melo and Weikum, 2009; Navigli and Ponzetto, 2010).

3. WOLF and sloWNet

Previous work on the development of sloWNet and WOLF (Erjavec and Fišer, 2006; Fišer and Sagot, 2008) has focused on benefitting from available resources of three different types: general and domain-specific bilingual dictionaries, parallel corpora and Wiki resources (Wikipedia and Wiktionaries).

More precisely, the development of the initial versions of WOLF and sloWNet was achieved in a three-step process. First, baseline versions of these wordnets were created (Fišer and Sagot, 2008) by using only *(literal,*

synset) pairs obtained from a word-aligned multilingual parallel corpus, which could be disambiguated based on all languages but French and Slovene, as well as pairs extracted from lexical resources (dictionaries, lexica and Wikipedia) via monosemous English literals: such pairs required no disambiguation. The resulting wordnets were relatively reliable but did not use full potential of the available lexical resources.

This is why we describe here a large-scale extension process, aiming at taking full advantage of these lexical resources for improving the coverage of both wordnets without lowering their accuracy. In the next section, we describe the lexical resources we used and how we extracted *(literal, synset)* candidates from them. In Section 5 we introduce the maximum entropy classifier we used for filtering these pairs and extending both wordnets, which are then evaluated in Section 6.

4. Extracting bilingual lexicons

In this section we describe the extraction of translation pairs from two types of resources: structured (general and domain-specific dictionaries and lexica), and semi-structured (Wikipedia articles). In the extraction process, our task is to extract as many translation variants for each word as possible in order to capture as many senses of that word as possible, in order to create *wordnet candidates* from the extracted translation pairs in the form of *(literal, synset) pairs*, i.e. translation of a source word with an assigned synset id from wordnet.

We used English, French and Slovene **Wiktionary** and extracted translation pairs for all parts-of-speech from these three resources on the basis of translation sections within the articles. The number of pairs extracted from each resource is given in Table 1. In order to extract the general vocabulary that was available in Wiktionary for French but not for Slovene we used a **traditional English-Slovene** (Grad *et al.*, 1999) **and a Slovene-English dictionary** (Grad and Leeming, 1999).

For domain-specific vocabulary we used **Wikispecies,** a taxonomy of living species that includes both Latin standard names and vernacular terms. We also obtained translation pairs from the domain-specific thesaurus **Eurovoc,** an on-line dictionary of informatics **islovar** and a **military glossary** (Korošec *et al.*, 2001) that will at least partially make up for the difference in the sizes of French and Slovene Wikispecies (cf. Table 1).

The result of our extraction process is two large bilingual lexicons containing all English-French and English-Slovene translation pairs with the name of the resources they originate from. The figures for both extracted bilingual lexicons are summarized in Table 1.

Input Resource	**En-Fr unique pairs**	**En-Sl unique pairs**
English wiktionary	39,286	6,052
French / Slovene wiktionary	59,659	7,029
Wikispecies	48,046	2,360
Slovene-English dictionary	–	72,954
English-Slovene dictionary	–	207,972
Eurovoc + specialized voc.	–	31,702
Total (duplicates removed)	**130,601**	**282,789**

Table 1: Bilingual lexicon extracted from structured resources

Less structured than dictionaries but still with a much more predefined structure than free text is the on-line multilingual collaborative encyclopaedia **Wikipedia.** We used English, French and Slovene Wikipedia for extracting bilingual lexicons by following inter-language links that relate two articles on the same topic in the two corresponding wikipedias. We enhanced the extraction process with a simple analysis of article bodies with which we resolved ambiguities arising from to the capitalization of article titles (e.g. *Grass-author, Grass-plant*). With the analysis we also identified synonyms for the key terms (e.g. *Cannabis*, also known as *marijuana*), their definitions (e.g. *Hockey* is a family of sports in which two teams play against each other by trying to manoeuvre a ball or a puck into the opponent's goal using a hockey stick.) and usage examples. The number of translation pairs extracted is shown in Table 2.

Input Resource	**En-Fr unique pairs**	**En-Sl unique pairs**
Wikipedia	286,818	32,161

Table 2: Bilingual lexicon extracted from Wikipedia

The bilingual entries we extracted from lexical resources are numerous. However, they suffer from an important drawback: they do not necessarily contain any additional information that can help to map them to PWN, neither do they contain contextual information from specific corpus occurrences. For example, an English-French entry we extracted from Wiktionary *(dog, chien)* does not contain any information that would make it possible for us to determine which of the 8 synsets containing *dog* as a literal in PWN would be appropriate to be translated with *chien* in WOLF. In Wiktionary articles, translations of a given word are sometimes organized by senses that are associated with short glosses. These have been compared to PWN glosses in order to map Wiktionary senses to PWN synsets (Bernhard and Gurevych, 2009). The first sentence of a Wikipedia article can be used in a similar way (Ruiz-Casado *et al.*, 2005). However, this is not the case for all Wiktionary entries or for other resources. Therefore, at this point, we assign to each translation pair *all* possible synset ids.

5. Automatic wordnet extension

5.1. Baseline wordnets

The first step in the development of sloWNet and WOLF was achieved in 2008, when the first versions were created (Fišer and Sagot, 2008). Then, all PWN literals were used for adding target language literals in the synsets found by the alignment-based approach, but only candidates generated from lexical resources via monosemous PWN literals were used.

After the restricted versions of wordnets were produced, they underwent some improvement steps that were performed independently, according to the specific needs of the two research teams. Quantitative information about the first two versions of WOLF are given in the second column in Table 3, which is compared to PWN 2.0, to the result of the work presented here (WOLF 0.2), as well as wordnets for

French that were developed by other researchers (French EuroWordNet (Vossen, 1999) and JAWS (Mouton and de Chalendar, 2010)). Parallel figures for sloWNet are shown as well. Table 3 uses the three Base Concept Sets (BCS) for assessing the coverage of the most basic word senses. BCS were introduced in the BalkaNet project and cover the 8,516 most basic synsets in wordnet.

	PWN 2.0	**WOLF 0.1.4**	**WOLF 0.1.6**	**WOLF 0.2**	**FWN**	**JAWS**
All	115,424	32,351	32,550	**46,449**	22,121	34,367
BCS1	1,218	869	870	**1,067**	1,211	760
BCS2	3,471	1,665	1,668	**2,519**	3,022	1,729
BCS3	3,827	1,796	1,801	**2,585**	2,304	1,706
Non-BCS	106,908	27,492	28,211	**40,278**	15,584	30,172
N	79,689	28,187	28,559	**36,933**	17,381	34,367
V	13,508	1,546	1,554	4,105	**4,740**	0
Adj	18,563	1,422	1,562	**4,282**	0	0
Adv	3,664	667	871	**1,125**	0	0

	PWN 2.0	**sloWNet 2.0**	**sloWNet 2.2**	**sloWNet 3.0**
All	115,424	29,108	17,817	**42,919**
BCS1	1,218	714	*1,203*	***1,208***
BCS2	3,471	1,361	*2,192*	***3,111***
BCS3	3,827	1,611	*1,232*	***2,698***
Non-BCS	106,908	25,422	*13,190*	***35,902***
N	79,689	22,927	16,234	**30,911**
V	13,508	1,547	1,097	**5,337**
Adj	18,563	4,376	429	**6,218**
Adv	3,664	258	57	**453**

Table 3: Quantitative data about the different versions of WOLF and sloWNet, and comparison with the PWN 2.0, the French wordnet from the EuroWordNet project (FWN) and the JAWS nominal wordnet for French.

5.2. Large-scale wordnet extension

Restricting the use of bilingual lexicon to monosemous English literals is a safe but limited approach that does not exploit the available resources to their full potential. However, using lexicon-based candidates generated from polysemous English literals is only possible if we can establish the likelihood with which a word should be added to a particular synset, i.e. can compute the semantic distance between a given French or Slovene literal and synset id. We designed such a technique based on already-existing French and Slovene wordnets, which we introduce in this Section.

Our technique relies on a probabilistic classifier that uses various features associated with each *(literal, synset)* candidate. The underlying idea is as follows: we start from baseline wordnets and a large set of lexicon-based candidates to be evaluated. We extract all *(literal, synset)* pairs that are already in the baseline wordnets and consider these candidates as valid ones (score 1) while all other candidates are considered invalid (score 0), thus creating a "copper standard", i.e., a reasonable although noisy training set for a probabilistic model. It is noisy for two reasons: first, our baseline wordnets do contain noise, as not all synsets were manually validated; second, and more importantly, many of our new candidates are valid even though they are not in the baseline wordnets. In fact, such candidates are exactly those we are looking for. In order to use the copper standard as training set for a classifier and then to assign scores to all candidates, we have to extract from them suitable features.

The most important feature that we use models the **semantic proximity** between a literal and a synset. Let us illustrate it on our running example *(dog, chien)*. In PWN, 8 synsets contain the literal *dog*, which is why we generated 8 different *(literal, synset)* candidates from this bilingual entry. We now need to know which of them are valid, i.e., to which of the 8 corresponding synsets the French literal *chien* should be added in WOLF. We therefore compute the semantic similarity of the literal *chien* w.r.t. each of these 8 synsets. For doing this, we first represent each WOLF synset by a bag of words obtained by extracting all literals from this synset and all the synsets up to 2 nodes apart in WOLF. For example, the synset *{andiron, firedog, dog, dog-iron}* in PWN, which is empty in WOLF 0.1.4, is represented by the bag of words *{appareil, mécanisme, barre, rayon, support, balustre, dispositif,…} (~device, mechanism, bar, shelve, baluster, device,…)*. Next, we use a distributional semantic model for evaluating the semantic similarity of *chien* w.r.t. this bag of words. We use the freely available SemanticVectors package (Widdows and Ferraro, 2008). The documents we used for building our distributional semantic models are 65,000 lemmatised webpages from the web-based frWaC corpus (Ferraresi *et al.*, 2010) for French (390,000 distinct lemmas) and 334,000 lemmatised paragraphs from the reference FIDA-Plus corpus (Arhar and Gorjanc, 2007) for Slovene (180,000 distinct lemmas). On our example, the semantic similarity between *chien* and the synset *{andiron, firedog, dog, dog-iron}* is only 0.035, while the similarity between *chien* and one of its valid synsets, *{dog, domestic dog, Canis familiaris}* is as high as 0.331.

Apart from that semantic similarity measure, we used several other features. Let us consider a candidate *(T, S)* that has been generated because our bilingual resources provided us with entries of the form $(E_1, T) \ldots (E_n, T)$, where all PWN literals E_i's are among S's literals. **The number of such PWN literals** is one of the features. **Each possible source** (e.g., the English Wiktionary) corresponds to one feature, which receives the value 1 if and only if at least one of the (E_i, T) entries was extracted from this source. Moreover we extract **the lowest polysemy index** among all E_i's: if one of the E_i's is monosemous, this feature receives the value 1; if the least polysemous E_i is in two PWN synsets, this features receives the value 2. The idea is that if the candidate is generated from at least one monosemous PWN literal, then it is very likely to be correct, whereas if it was generated from only highly polysemous PWN literals, it is much more questionable. Finally, **the number of tokens** in T is used as a feature (literals with many tokens are usually not translations of PWN literals but rather glosses, and are therefore incorrect).

Based on these features, we trained one classifier per language using the Maximum-Entropy package *megam* (Hal Daumé III, 2004). A look at the resulting models (not shown here for space restrictions) shows that the

semantic similarity we computed is relevant as it is the feature with the highest weight. As expected, the lowest polysemy index among English literals also contributes positively, as does the number of different English literals yielding the generation of the candidate, and the number of sources involved. On the other hand, also as expected, the number of tokens in the target language literal negatively contributes to the certainty score.

The result of our classifiers on a given *(literal, synset)* candidate is a score between 0 (bad candidate) and 1 (good candidate). We empirically set the threshold at 0.1 (see Section 6.1) for further addition in the corresponding wordnet. This resulted in retaining 55,159 French candidates (out of 177,980) and 68,070 Slovene candidates (out of 685,633). Among the 55,159 French candidates, 15,313 (28%) correspond to *(literal, synset)* pairs already present in WOLF 0.1.6, which means that 39,823 (72%) new ones were added. As a consequence, 13,899 synsets that were empty in WOLF 0.1.6 now have at least one French literal. Among the 68,070 Slovene candidates, 5,056 (7%) correspond to *(literal, synset)* pairs already present in sloWNet, which means that 63,010 (63%) new ones were added; as a consequence, 25,102 synsets that were empty in sloWNet have now at least one Slovene literal.

Quantitative information on the resulting wordnets (WOLF 0.2 and sloWNet 3.0) is provided in Table 3. In short, WOLF 0.2.0 has 43% more non-empty synsets than before the extension, and sloWNet 3.0 as much as 141% more. For *(literal, synset)* pairs, the increase is even higher: the extension of WOLF has increased the number of such pairs from 46,411 to 76,436 (+65%), and the extension of sloWNet has increased this number from 24,081 to 82,721 (+244%).

6. Evaluation

Before assessing the accuracy of the wordnets extended using the above-described approach, we begin with a manual evaluation of the extension step *per se*. We evaluate the accuracy of the candidates we obtained as well as the accuracy of the candidates we discarded. Next, we perform an automatic evaluation of the extended wordnets against a gold standard or a comparable resource developed by other authors.

6.1. Manual evaluation of wordnet extension

For measuring the accuracy of our extension approach (see Section 4.3), we randomly selected 400 *(literal, synset)* candidates for each language and evaluated them manually, using only two tags: "OK" if it would be correct to add that literal to the synset, and "NO" if it would be wrong, regardless of the cause of the error and how semantically close it was to the synset. The accuracy of a set of candidates is the proportion of candidates tagged "OK". Moreover, in order to assess the quality of our scoring technique, we compared the accuracy of the candidates per quartile w.r.t. their certainty scores. The results (see Table 4) show a strong correlation between the certainty score and the accuracy of the candidates, leading us to set the threshold value at 0.1. Other threshold values could have been used: higher values would have provided candidates with even a higher accuracy but the scale of the wordnet extension would have been lower; on the other hand, lower threshold values would have extended our wordnets even more, but would have introduced much more noise. The 0.1 value, which corresponds approximately to the upper quartile for French and to the upper decile for Slovene, seemed to provide a good balance — even if the candidates retained exhibit a higher precision in French (83%) than in Slovene (64%), because of many rare and archaic words coming from the English-Slovene dictionary that have often been evaluated as incorrect —, and lead us to retain similar numbers of candidates in both languages (55,159 for French and 68,070 for Slovene), even though many more Slovene candidates (685,633) were generated in comparison to French candidates (177,980).

Language	**French**	**Slovene**
No. of candidates evaluated manually	400	400
…among which are added to the wordnet (score ≥ 0.1)	**110 (27%)**	**36 (9 %)**
Accuracy over all candidates	52%	25 %
Accuracy of candidates added to the wordnet (score ≥ 0.1)	**81%**	**64 %**
Acc. of discarded candidates (score < 0.1)	40%	21 %
Accuracy in the upper (4th) quartile	83%	44 %
Accuracy in the third quartile	63%	32 %
Accuracy in the second quartile	41%	13 %
Accuracy in the lower (1st) quartile	20%	10 %

Table 4: Manual evaluation of *(literal, synset)* candidates generated for extending WOLF and sloWNet.

6.2. Evaluation against other wordnets

In this section we report the results of automatic evaluation of the generated wordnets, which are compared to other wordnets that exist for the two target languages. However, such an evaluation is only partial, because the detected discrepancies between the two resources are not only errors in our automatically created wordnets but can also stem from a missing literal in the resource it is compared to. Automatic evaluation was performed on non-empty synsets, which means that adjectival and adverbial synsets in WOLF could not be evaluated this way at all because other existing French wordnets do not cover them.

The results of automatic evaluation of WOLF w.r.t. the FWN are not given here in full details due to space restrictions. However, let us consider first non-empty synsets in FWN. Any *(literal, synset)* pair that is common to both resources is considered correct. There also exist WOLF pairs that are not present in FWN although their synsets are not empty in FWN. Some of these are correct (i.e. even non-empty FWN synsets are incomplete), others not (i.e. these are errors in WOLF). For an estimate of the precision of such pairs, we manually evaluated 100 randomly selected such nominal and 100 verbal pairs. This allowed us to estimate the number of correct vs. incorrect *(literal, synset)* pairs among those present in WOLF but not in FWN. And last but not least, many synsets are empty in FWN. We manually validated 100 randomly chosen pairs for such synsets and obtained an accuracy as high as 92%. In fact, empty synsets in FWN are rare or specific concepts whose literals are often monosemous and therefore easy to translate. Adding up the (exact or

estimated) number of valid *(literal, synset)* pairs of all types, we reach a total of ~65,690 valid pairs out of 76,436, hence a **~86% accuracy**.

The reference we used for evaluating sloWNet is a manually built gold standard (SWN) that was built by validating the results of the preliminary Slovene wordnet construction experiments based on the Serbian wordnet (Fišer, 2008). The result is approximately the same: we get a **~85% accuracy**.

7. Conclusion

We have described the different resources and techniques we used for extending WOLF and sloWNet, wordnets for French and Slovene that were built automatically. By using various features including distributional similarity, we were able to reuse automatically extracted bilingual lexicons for translating and disambiguating polysemous literals, which had been dealt only with word-aligned corpora for building the first versions of these wordnets. The result of our work is freely available lexical semantic resources that are large and accurate enough for being used in real NLP applications.

The extended wordnets, which are much larger than the previous versions, were then carefully evaluated in terms of accuracy. The accuracy of *(literal, synset)* pairs is estimated at 86% for WOLF 0.2 and 85% for sloWNet 3.0. These figures show that both resources have a much higher coverage than the baseline wordnets and that they outperform the French EuroWordNet as well as JAWS, that currently only covers French nouns. A direct comparison with other related resources developed by Navigli and Ponzetto (2010) and di Melo and Weikum (2010) is not straightforward because even though the resources we used overlap to a great extent, their aim was to create a multilingual network while we focused only on the two target languages. Also, while Navigli and Ponzetto (2010) machine-translated the missing translations, we only use resources that were created by humans, which is why we expect to have more accurate translations but would have to carry out a detailed comparison to be certain. While di Melo and Weikum's (2010) wordnet for French has a slightly higher accuracy, it is smaller than ours. This shows that the approach we used, namely trying to benefit as much as possible from available resources using basic NLP tools only, is very efficient for building large-scale reliable wordnets.

References

Arhar, Š. and Gorjanc, V. (2007). Korpus FidaPLUS: nova generacija slovenskega referenčnega korpusa. *Jezik in slovstvo*, **52**(2).

Bernhard, D. and Gurevych, I. (2009). Combining Lexical Semantic Resources with Question & Answer Archives for Translation-Based Answer Finding. *Proc. of ACL-IJCNLP'09*.

Daumé III, H. (2004). *Notes on CG and LM-BFGS optimization of logistic regression*.

Diab, M. (2004). The Feasibility of Bootstrapping an Arabic WordNet leveraging Parallel Corpora and an English WordNet. In *Proc. of NEMLAR*, Cairo, Egypt.

Dyvik, H. (2004). Translations as semantic mirrors: from parallel corpus to wordnet. In *Language and Computers* **49**(1), p. 311-326.

Erjavec, T. and Fišer, D. (2006). Building the Slovene Wordnet : first steps, first problems. In *Proc. of the 3rd Int. WordNet Conference*, Jeju, Korea.

Fellbaum, C., editor. 1998. *WordNet: An Electronic Database*. MIT Press, Cambridge, USA.

Ferraresi, A., Bernardini, S., Picci, G. and Baroni, M. (2010). Web Corpora for Bilingual Lexicography: A Pilot Study of English/French Collocation Extraction and Translation. In Xiao, R. (ed.) *Using Corpora in Contrastive and Translation Studies*. Newcastle: Cambridge Scholars Publishing.

Fišer, D. and Benoît, S. (2008). Combining multiple resources to build reliable wordnets. In *Proc. of TSD'08*, Brno, Czech Republic.

Fung, P. (1995). A pattern matching method for finding noun and proper noun translations from noisy parallel corpora. In *Proc. of ACL'95*.

Grad, A. and Leeming, H. (1999). *Slovene-English Dictionary*, DZS

Grad, A., Škerlj, R. and Vitorovič, N. (1999). *English-Slovene Dictionary*, DZS.

Ide, N., Erjavec, T., Tufiş, D. (2002). Sense Discrimination with Parallel Corpora. In *Proc. of the ACL'02 Workshop on Word Sense Disambiguation*, Philadelphia, USA.

Knight, K. and Luk, S. K. (1994). Building a large-scale knowledge base for machine translation. In *Proc. of AAAI'94*.

Korošec, T., Fekonja, M., Jehart, A., Pečelin, F., Ulčar, M. and others (2002). *Vojaški slovar*. Ljubljana, Slovenia: Ministrstvo za obrambo.

Krstev, C., Pavlović-Lažetić, G., Vitas, D. and Obradović, I. (2004). Using textual resources in developing Serbian WordNet. In *Romanian Journal of Information Science and Technology*. Dan Tufiş (ed.), **7**(1–2), pp. 147-161.

de Melo, G. and Weikum, G. (2009). Towards a universal wordnet by learning from combined evidence. In *Proc. of CIKM'09*.

Mouton, C. and de Chalendar, G. (2010). JAWS: Just Another WordNet Subset. In *Proc. of TALN'10*, Montreal, Canada.

Navigli, R., Ponzetto, S. P. (2010). BabelNet: Building a Very Large Multilingual Semantic Network. In *Proc of. ACL'10*, Uppsala, Sweden.

Ponzetto, S. P. and Navigli, R. (2009). Large-scale taxonomy mapping for restructuring and integrating Wikipedia. In *Proc. of IJCAI'09*.

Ruiz-Casado, M., Alfonseca, E. and Castells, P. (2005). Automatic Assignment of Wikipedia Encyclopedic Entries to WordNet Synsets. In *Proc. of AWIC'2005*. pp. 380–386

Tufiş, D. (2000). BalkaNet — Design and Development of a Multilingual Balkan WordNet. In *Romanian Journal of Information Science and Technology Special Issue*, **7**(1-2).

Vossen, P., editor. (1998). *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Kluwer, Dordrecht, The Netherlands.

Widdows, D. and Ferraro, K. (2008). Semantic vectors: a scalable open source package and online technology management application. In *Proc. of LREC'08*, Marrakech, Morocco.