# Reserating the awesometastic:
# An automatic extension of the WordNet taxonomy for novel terms

**David Jurgens**
School of Computer Science
McGill University
jurgens@cs.mcgill.ca

**Mohammad Taher Pilehvar**
Department of Computer Science
Sapienza University of Rome
pilehvar@di.uniroma1.it

## Abstract

This paper presents CROWN, an automatically constructed extension of WordNet that augments its taxonomy with novel lemmas from Wiktionary. CROWN fills the important gap in WordNet's lexicon for slang, technical, and rare lemmas, and more than doubles its current size. In two evaluations, we demonstrate that the construction procedure is accurate and has a significant impact on a WordNet-based algorithm encountering novel lemmas.

## 1 Introduction

Semantic knowledge bases are an essential, enabling component of many NLP applications. A notable example is WordNet (Fellbaum, 1998), which encodes a taxonomy of concepts and semantic relations between them. As a result, WordNet has enabled a wide variety of NLP techniques such as Word Sense Disambiguation (Agirre et al., 2014), information retrieval (Varelas et al., 2005), semantic similarity (Pedersen et al., 2004; Bär et al., 2013), and sentiment analysis (Baccianella et al., 2010). However, semantic knowledge bases such as WordNet are expensive to produce; as a result, their scope and domain are often constrained by the resources available and may omit highly-specific concepts or lemmas, as well as new terminology that emerges after their construction. For example, WordNet does not contain the nouns "stepmom," "broadband," and "prequel."

Because of the coverage limitations of WordNet, several approaches have attempted to enrich WordNet with new relations and concepts. One group of approaches has enriched WordNet by aligning its structure with that of other resources such as Wikipedia or Wiktionary (Ruiz-Casado et al., 2005; Navigli and Ponzetto, 2012; Miller and Gurevych, 2014; Pilehvar and Navigli, 2014). However, because these approaches identify corresponding lemmas with identical lexicalizations, they are often unable to directly add novel lemmas to the existing taxonomic structure. The second group of approaches performs taxonomy induction to learn hypernymy relation-

ships between words (Moro and Navigli, 2012; Meyer and Gurevych, 2012). However, these approaches often produce separate taxonomies from WordNet, which are also generally not readily accessible as resources.

We introduce a new resource CROWN (**C**ommunity-en**R**iched **O**pen **W**ord**N**et) that extends the existing WordNet taxonomy, more than doubling the existing number of synsets, and attaches these novel synsets to their appropriate hypernyms in WordNet. Novel sense data is extracted from Wiktionary, a large-scale collaboratively-constructed dictionary, and attached using multiple heuristics. CROWN fills an important gap in WordNet's limited coverage of both domain-specific lemmas and slang terminology and idioms.[1] In two experiments, we demonstrate that (1) our construction process accurately associates a novel sense with its correct hypernym and (2) the resulting resource has an immediate benefit for existing WordNet-based applications. Importantly, CROWN v1.0 is publicly available and released in WordNet format, making it seamlessly integratable with all existing WordNet libraries and tools.

## 2 Wiktionary

Wiktionary is a multilingual online dictionary that, as of May 2014, contains more than 470K English gloss definitions. Thanks to its collaboratively-constructed nature, Wiktionary provides a high coverage of novel domain-specific, idiomatic and slang terms or meanings, across all parts of speech, while featuring a wide variety of linguistic information such as morphology, etymology, pronunciation and alternative lexicalizations of a lemma. Given these characteristics, Wiktionary is an ideal resource for improving the coverage of hand-crafted lexicons, such as WordNet.

In addition to definitions, Wiktionary contains two sources of semantic relations. First, the Wiktionary entry

---

[1] For example, "reserate" is correctly included in CROWN as a hypernym of `unlock%2:35:00::` (to open the lock of) and "awesometastic" as a synonym of `fantastic%3:00:00:extraordinary:00` (extraordinarily good or great).

for a lemma may contain a note stating its relationship with another lemma. Second, Wiktionary includes a separate thesaurus, Wikisaurus, which specifies (1) a lemma and its gloss and (2) all other lemmas sharing a relation with that sense. However, these Wiktionary relations cannot directly be used to enrich WordNet for two reasons. First, Wiktionary entries are defined in terms of lemmas, rather than senses. As a result, directly ontologizing the resource or integrating its semantic relations requires disambiguating each relation's lemmas, which is not always possible due to the limited context. Second, semantic relations in Wiktionary are infrequent, with 19.8% of all words having any specified relation and only 0.3% having a hypernym relation. As a result of this sparsity, structure alignment-based approaches for extending WordNet cannot be directly applied.

# 3 Extending WordNet

CROWN is created by identifying lemmas that are out of vocabulary (OOV) in WordNet but have one or more associated glosses in Wiktionary. A new synset is created for that lemma and a hypernym relation is added to the appropriate WordNet synset. The CROWN attachment process rates hypernym candidates using two methods. First, where possible, we exploit structural or morphological information to identify highly-probable candidates. Second, following previous work on resource alignment showing that lexical overlap accurately measures gloss semantic similarity (Meyer and Gurevych, 2011; Navigli and Ponzetto, 2012), candidates are found by measuring the similarity of the Wiktionary gloss with the glosses of synsets found by a constrained search of the WordNet graph. We note that attaching OOV lemmas by first aligning WordNet and Wiktionary is not possible due to relation sparsity within Wiktionary, where most OOV words would not be connected to the aligned network. Following, we first describe the Wiktionary preprocessing steps and then detail both OOV attachment methods.

## 3.1 Preprocessing

Wiktionary was parsed using JWKTL (Zesch et al., 2008) to extract the text associated with each Wiktionary definition and remove Wiktionary markup. The extracted texts were then partitioned into two sets: (1) those expressing a lexicalization, e.g., "1337" is an alternative spelling of "elite" and (2) those indicating a definition. Novel lexicalizations that are not already handled by the WordNet morphological analyzer (Morphy) were added to the lexicalization exception lists in CROWN.

Definitions are processed using two methods to identify a set of candidate lemmas whose senses might be identical or near to the appropriate hypernym synset. First, candidates are obtained by parsing the gloss with Stanford CoreNLP (Manning et al., 2014) and extract-

ing the head word and all other words joined to it by a conjunction. Second, additional candidates are collected from the first hyperlinked term or phrase in the gloss, which is similar to the approach of Navigli and Velardi (2010) for hypernym extraction in Wikipedia. Candidates are then filtered to ensure that (1) they have the same part of speech as the definition's term and (2) they are defined in WordNet, which is necessary for the attachment.

## 3.2 Structural and Lexical Attachment

Three types of structural or lexical heuristics were used to attach OOV lemmas when the appropriate data was available. First, Wikisaurus or Wiktionary synonym relations create sets of mutually-synonymous lemmas, which may contain OOV lemmas. The common hypernym of these lemmas is estimated by computing the most frequent hypernym synset for all the senses of the set's lemmas that are in WordNet. Any OOV lemma also in the set is then attached to this estimated hypernym.

Second, some Wiktionary glosses follow regular patterns that identify a particular meaning. Two pattern heuristics were used: (1) a group of *Person* patterns and (2) a *Genus* pattern. The *Person* patterns match glosses that start with phrases such as "somebody who." Senses with such glosses have their set of candidate attachments restricted to descendants of the human sense of the noun *person*; the sense is then attached to a descendant using the gloss ranking procedure for lexical attachment (described below). The *Genus* pattern matches glosses that start with "Any member of the" and later contain a proper noun matching a scientific genus in WordNet; in such cases the OOV lemma is attached to the same hypernym as the synsets with a holonymy relation to the genus's synset.

Third, an *Antonymy* heuristic is used to identify OOV lemmas with an antonym relation to lemmas already in WordNet. OOV lemmas are tested for having a prefix indicating it could be an antonym, e.g., "anti." If the lemma formed from the remainder after prefix is in WordNet, then the OOV lemma is treated as its antonym and attached to the antonym's hypernym. Furthermore, the two synsets are marked as antonyms in CROWN.

## 3.3 Gloss-based Attachment

Each OOV lemma is associated with one or more Wiktionary senses, $s_{1...n}$, where each sense $s_i$ is associated with a set of lemmas $l_i$, one of whose senses may be its hypernym. The gloss-based attachment method analyzes each sense separately, first generating a set of candidate hypernym synsets and then ranking each synset according to its gloss similarity, both defined next. Ultimately the OOV lemma is attached to the highest-scoring synset across all of its Wiktionary senses. This procedure is intended to maximize precision by attaching only the

| | | | |
|---|---|
| **ukWaC** | microsoft, e-learning, helpline, mp3, unsubscribe |
| **Twitter** | selfie, retweet, hella, bday, homie |
| **Wikipedia** | admin, verifiability, bio, sockpuppetry, same-sex |

Table 1: Examples of high-frequency lemmas in CROWN but not in WordNet, from three corpora.

lemma's dominant sense, though we note that most OOV lemmas are monosemous.

The initial set $C$ of candidate hypernym synsets for Wiktionary sense $s_i$ is generated from the union of the synsets of the lemmas in $l_i$. Then, $C$ is expanded by including all WordNet synsets reachable from each synset $c_i \in C$ by a path of hypernym or hyponym edges, where a path (1) has at most three edges and (2) contains at most one hypernym edge. The second constraint is designed to avoid including overly-general concepts.

The glosses of the synsets in $C$ are then compared with the Wiktionary sense's gloss. Directly comparing glosses with string similarity measures omits the important detail that certain lemmas can be highly-specific and most strongly indicate that two glosses refer to the same concept. Therefore, prior to comparison, the lemmas occurring in all glosses are assigned a weight $-\log\frac{1}{f(w)}$, where $f(w)$ denotes the number of glosses in which lemma $w$ appeared. Glosses' similarity is measured by (1) lemmatizing their texts and computing the lemmas in common, and then (2) summing the weights of the in-common lemmas. This similarity function assigns higher scores to glosses sharing more specific concepts.

### 3.4 Resource Creation

The resulting attachments are converted into WordNet lexicography files and then integrated with the existing WordNet taxonomy using the GRIND program. Table 2 shows the resulting statistics for CROWN in comparison to WordNet. The attachment process more than doubles the number of synsets and adds a significant number of new lexicalizations which are essential for capturing common spelling variants that are not reflected in WordNet. Additionally, 4739 new antonym relations were added. Of the OOV lemmas, 87.8% were attached using the lexical attachment procedure. Of the remaining, the *Person* and *Antonymy* heuristics were the most frequently used, accounting for 4.2% and 2.7% of cases respectively. The infrequent use of the structural and lexical heuristics underscores the sparsity of the available data in Wiktionary for straight-forward attachments.

As an initial test of additional content present in CROWN but not in WordNet, all lemmas unique to CROWN were extracted and their occurrences counted in three corpora: (1) all of the English Wikipedia, (2) the web-gathered ukWaC corpus (Ferraresi et al., 2008), and (3) a sample of 50M microtext message from Twit-

| PoS | WordNet synsets | new CROWN synsets | new CROWN lex. variants |
|---|---|---|---|
| Noun | 82115 | 124967 | 29563 |
| Verb | 13767 | 16199 | 43318 |
| Adj. | 18156 | 25534 | 6902 |
| Adv. | 3621 | 2031 | 481 |

Table 2: The number of synsets in WordNet and new synsets and lexicalizations added by CROWN.

ter. Table 1 shows five example high-frequency lemmas from each corpus that are only present in CROWN , highlighting the types of commonly-recognized terms not in WordNet due to their technical, informal, or recently-created nature. Indeed, "selfie" was only recently included in the Merriam Webster dictionary as of 2014,[2] demonstrating the potential for quickly integrating new terminology into CROWN from the frequently-updated entries of Wiktionary.

## 4 Evaluation

Two evaluations were performed. The first estimates attachment accuracy by simulating OOV attachment with lemmas that are already in WordNet. The second calculates the benefit of using CROWN in an example application using a WordNet-based algorithm to measure similarity.

### 4.1 WordNet Replication

No standard dataset exists for where OOV lemmas should be attached to WordNet; therefore in the first evaluation, we assess construction accuracy by simulating the inclusion of OOV lemmas using those already in WordNet, which allows testing on tens of thousands of lemmas. Specifically, the CROWN attachment approach is used to reattach all monosemous lemmas in WordNet. We opted for monosemous terms as they can have only one valid location in the taxonomy.

### 4.1.1 Methodology

Glosses were extracted for 36,605 of the 101,863 nouns that were monosemous in WordNet and also present in Wiktionary, and for 4668 of the 6277 verbs matching the same condition. These glosses were then provided as input to the CROWN attachment process. We note that these lemmas are not necessarily monosemous in Wiktionary, with nouns and verbs having on average 1.40 and 1.76 senses, respectively; however, the construction process will attach only the highest-scoring of these senses. Once a lemma is attached, accuracy is measured as the number of hyponym or hypernym edges away that CROWN placed the lemma from its original position.

---

[2]http://www.merriam-webster.com/new-words/2014-update.htm

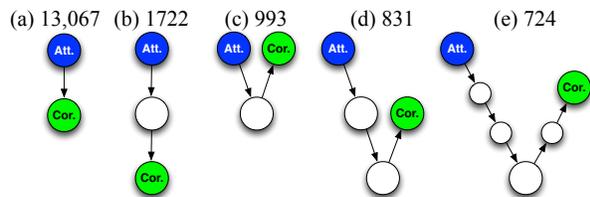(a) 13,067 (b) 1722 (c) 993 (d) 831 (e) 724

Figure 1: The five most-frequent error patterns and their frequencies seen in the results of monosemous lemma evaluation. Graphs show the attachment point (Att.) and correct hypernym synset (Cor.), with downward edges indicating hypernym relations and upward indicating hyponym. The overall error trend reveals that the vast majority of error was due to attaching a new sense to a more-specific concept than its actual hypernym.

### 4.1.2 Results

The CROWN construction process was able to attach 34,911 of the 36,605 monosemous noun lemmas (95.4%) and 4209 of the 4668 verb lemmas (90.2%). The median error for attaching monosemous nouns was three edges and for verbs was only one edge, indicating the attachment process is highly accurate for both. The most common form of error was attaching the OOV lemma to a hyponym of the correct hypernym, occurring in 13,067 of the erroneous attachments.

Figure 1 shows the five most common displacement patterns when incorrectly attaching a monosemous noun, revealing that the majority of incorrect placements were to a more-specific concept than what was actually the hypernym. Furthermore, examining the 50 furthest-away noun and verb placements, we find that 28% of nouns and 20% of verbs were attached using a novel sense of the lemma not in WordNet (but in Wiktionary) and the placement is in fact reasonable. As a result, the median error is likely an overestimate of the expected error for the CROWN construction process.

### 4.2 Application-based evaluation

Semantic similarity is one of the core features of many NLP applications. The second evaluation measures the performance improvement of using CROWN instead of WordNet for measuring semantic similarity when faced with slang or OOV lemmas. Notably, prior semantic similarity benchmarks such as SimLex-999 (Hill et al., 2014) and the ESL test questions (Turney, 2001) have largely omitted these types of words. However, the recent dataset of SemEval-2014 Task 3 (Jurgens et al., 2014) includes similarity judgments between a WordNet sense and a word not defined in WordNet's vocabulary or with a slang interpretation not present in WordNet.

|  | All | Regular | OOV | Slang |
|---|---|---|---|---|
| WordNet | 0.195 | 0.463 | 0.0 | -0.170 |
| CROWN | 0.248 | 0.452 | 0.448 | 0.138 |
| GST Baseline | 0.148 | 0.283 | 0.148 | 0.018 |
| Best System | 0.389 | 0.529 | 0.501 | 0.146 |

Table 3: The Pearson correlation performance of ADW when using the WordNet and CROWN semantic networks on the word-to-sense test dataset of SemEval-2014 Task 3. We also show results for the string-based baseline system (GST) and for the best participating system in the word-to-sense comparison type of Task 3.

### 4.2.1 Methodology

Semantic similarity was measured using the similarity algorithm of Pilehvar et al. (2013), ADW,[3] which first represents a given linguistic item (such as a word or a concept) using random walks over the WordNet semantic network, where random walks are initialized from the synsets associated with that item. The similarity between two linguistic items is accordingly computed in terms of the similarity of their corresponding representations. ADW is an ideal candidate for measuring the impact of CROWN for two reasons. First, the algorithm obtains state-of-the-art performance on both word-based and sense-based benchmarks using only WordNet as a knowledge source. Second, the method is both unsupervised and requires no parameter tuning, removing potential performance differences between WordNet and CROWN being due to these factors.

To perform the second experiment, the ADW algorithm was used to generate similarity judgments for the data of Task 3, changing only the underlying semantic network to be either (1) the WordNet 3.0 network, with additional edges from disambiguated glosses,[4] or (2) the same network with novel synsets from CROWN. As the ADW algorithm is unchanged between settings, any performance change is due only to the differences between the two networks. Performance is measured using Pearson correlation with the gold standard judgments.

### 4.2.2 Results

Of the 60 OOV lemmas and 38 OOV slang terms in the test data, 51 and 26 were contained in CROWN, respectively. Table 3 shows the Pearson correlation performance of ADW in the two settings for all lemmas in the dataset, and for three subsets of the dataset: OOV, slang, and regular lemmas, the latter of which are in WordNet; the bottom rows show the performance of the Task's best participating system for the word-to-sense comparison type (Kashyap et al., 2014) and the most competi-

tive baseline, based on Greedy String Tiling (GST) (Wise, 1996).

ADW sees large performance improvements in the OOV and slang words when using CROWN instead of WordNet, which are both statistically significant at p<0.01. The overall improvement of ADW would place it as the fifth best system in this comparison type of Task 3. The performance on regular in-WordNet and OOV lemmas is approximately equal, indicating the high accuracy of OOV hypernym attachment in CROWN. Notably, on OOV and Slang, the unsupervised ADW, when coupled with the additional information in CROWN , produces competitive results with the best performing system, which is a multi-feature supervised system utilizing extensive external dictionaries and distributional methods.

## 5 Related Work

Most related is the work of Poprat et al. (2008), who attempted to automatically build an extension of WordNet with biomedical terminology; however, they were unsuccessful in constructing the resource. Other work has attempted to leverage distributional similarity techniques (Snow et al., 2006) or exploit the structured information in Wikipedia (Ruiz-Casado et al., 2005; Toral et al., 2008; Ponzetto and Navigli, 2009; Yamada et al., 2011) in order to extend WordNet with new synsets. However, structure-based approaches are limited only to the concepts appearing in Wikipedia article titles, which almost always correspond to noun concepts. Distributional and probabilistic approaches are also limited to OOV terms for which it is possible to gather enough statistics. As Wiktionary contains all parts of speech and our method is independent of word frequency, neither limitation applies to this work.

Other related work has attempted to tap resources such as Wikipedia for automatically constructing new ontologies (Suchanek et al., 2007; Dandala et al., 2012; Moro and Navigli, 2012; Meyer and Gurevych, 2012), extending existing ones through either alignment-based methods (Matuschek and Gurevych, 2013; Pilehvar and Navigli, 2014) or inferring the positions of new senses by their shared attributes which are extracted from text (Reisinger and Paşca, 2009). Extension and alignment approaches based on Wikipedia are limited mainly to noun concepts in Wikipedia; furthermore, these techniques cannot be directly applied to Wiktionary because its lack of taxonomic structure would prevent adding most OOV data to the existing WordNet taxonomy.

## 6 Conclusion

This work has introduced CROWN version 1.0, a new extension of WordNet that merges sense definitions from Wiktionary to add new hypernym and antonym relations. The resulting taxonomy has more than doubled the number of synsets in WordNet and includes many technical and slang terms, as well as non-standard lexicalizations. CROWN is released in the same format as WordNet[5] and therefore is fully compatible with all existing WordNet-based tools and libraries. Furthermore, the software for building CROWN has been opened-sourced and will be updated with future versions. In two experiments we demonstrated that the CROWN construction process is accurate and that the resulting resource has a real benefit to WordNet-based applications.

Immediate future work will add support for including new lemmas as synonyms in existing synsets and linking newly-created synsets with all appropriate types of WordNet semantic relationship. Longer-term future work will pursue more sophisticated methods for taxonomy enrichment to improve the quality of integrated content and will aim to integrate additional dictionaries, with a special emphasis on adding domain-specific terminology.

## References

Eneko Agirre, Oier Lopez de Lacalle, and Aitor Soroa. 2014. Random walks for knowledge-based Word Sense Disambiguation. *Computational Linguistics*, 40(1):57–84.

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC)*, volume 10, pages 2200–2204, Valletta, Malta.

Daniel Bär, Torsten Zesch, and Iryna Gurevych. 2013. DKPro Similarity: An open source framework for text similarity. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 121–126, Sofia, Bulgaria.

Bharath Dandala, Rada Mihalcea, and Razvan Bunescu. 2012. Towards building a multilingual semantic network: Identifying interlingual links in wikipedia. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics (*SEM)*, pages 30–37, Montreal, Canada.

Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Database*. MIT Press, Cambridge, MA.

Adriano Ferraresi, Eros Zanchetta, Marco Baroni, and Silvia Bernardini. 2008. Introducing and evaluating ukWaC, a very large web-derived corpus of English. In *Proceedings of the 4th Web as Corpus Workshop (WAC-4)*, Morocco.

Felix Hill, Roi Reichart, and Anna Korhonen. 2014. Simlex-999: Evaluating semantic models with (genuine) similarity estimation. *arXiv preprint arXiv:1408.3456*.

David Jurgens, Mohammad Taher Pilehvar, and Roberto Navigli. 2014. Semeval-2014 task 3: Cross-level semantic similarity. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval-2014)*, pages 17–26, Dublin, Ireland.

---

[5]Both the software for creating CROWN and the data itself are available at `https://github.com/davidjurgens/crown`.

Abhay Kashyap, Lushan Han, Roberto Yus, Jennifer Sleeman, Taneeya Satyapanich, Sunil Gandhi, and Tim Finin. 2014. Meerkat mafia: Multilingual and cross-level semantic textual similarity systems. In *Proceedings of the 8th International Workshop on Semantic Evaluation*, pages 416–423, Dublin, Ireland.

Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 55–60, Baltimore, Maryland.

Michael Matuschek and Iryna Gurevych. 2013. Dijkstra-WSA: A graph-based approach to word sense alignment. *Transactions of the Association for Computational Linguistics (TACL)*, 1:151–164.

Christian M. Meyer and Iryna Gurevych. 2011. What psycholinguists know about Chemistry: Aligning Wiktionary and WordNet for increased domain coverage. In *Proceedings of the 5th International Joint Conference on Natural Language Processing*, pages 883–892, Chiang Mai, Thailand.

Christian M. Meyer and Iryna Gurevych. 2012. OntoWiktionary constructing an ontology from the collaborative online dictionary Wiktionary. In *Semi-Automatic Ontology Development: Processes and Resources*, chapter 6, pages 131–161. IGI Global.

Tristan Miller and Iryna Gurevych. 2014. WordNet–Wikipedia–Wiktionary: Construction of a three-way alignment. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC)*, pages 2094–2100, Reykjavik, Iceland.

Andrea Moro and Roberto Navigli. 2012. WiSeNet: Building a Wikipedia-based semantic network with ontologized relations. In *Proceedings of the 21st ACM Conference on Information and Knowledge Management (CIKM)*, pages 1672–1676, Maui, HI, USA.

Roberto Navigli and Simone Paolo Ponzetto. 2012. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193:217–250.

Roberto Navigli and Paola Velardi. 2010. Learning Word-Class Lattices for definition and hypernym extraction. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1318–1327.

Ted Pedersen, Siddharth Patwardhan, and Jason Michelizzi. 2004. WordNet:: Similarity: measuring the relatedness of concepts. In *Proceedings of Fifth Annual Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL)*, pages 38–41, Boston, Massachusetts.

Mohammad Taher Pilehvar and Roberto Navigli. 2014. A robust approach to aligning heterogeneous lexical resources. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014)*, pages 468–478, Baltimore, Maryland.

Mohammad Taher Pilehvar, David Jurgens, and Roberto Navigli. 2013. Align, Disambiguate and Walk: a Unified Approach for Measuring Semantic Similarity. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1341–1351, Sofia, Bulgaria.

Simone Paolo Ponzetto and Roberto Navigli. 2009. Large-scale taxonomy mapping for restructuring and integrating Wikipedia. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2083–2088, Pasadena, California, USA.

Michael Poprat, Elena Beisswanger, and Udo Hahn. 2008. Building a BioWordNet by using WordNet's data formats and WordNet's software infrastructure: a failure story. In *Proceedings of the Workshop on Software Engineering, Testing, and Quality Assurance for Natural Language Processing*, pages 31–39, Columbus, Ohio.

Joseph Reisinger and Marius Paşca. 2009. Latent variable models of concept-attribute attachment. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2-Volume 2*, pages 620–628, Suntec, Singapore.

Maria Ruiz-Casado, Enrique Alfonseca, and Pablo Castells. 2005. Automatic assignment of Wikipedia encyclopedic entries to WordNet synsets. In *Proceedings of the Third International Conference on Advances in Web Intelligence*, pages 380–386, Lodz, Poland.

Rion Snow, Daniel Jurafsky, and Andrew Y. Ng. 2006. Semantic Taxonomy Induction from Heterogenous Evidence. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics (COLING-ACL)*, pages 801–808, Sydney, Australia.

Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. 2007. YAGO: A core of semantic knowledge. unifying WordNet and Wikipedia. In *Proceedings of the 16th World Wide Web Conference (WWW)*, pages 697–706, Banff, Alberta, Canada.

Antonio Toral, Rafael Muoz, and Monica Monachini. 2008. Named Entity WordNet. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC)*, pages 741–747.

Peter Turney. 2001. Mining the web for synonyms: PMI-IR versus LSA on toefl. In *Proceedings of the Twelfth European Conference on Machine Learning (ECML-2001)*, pages 491–502, London, UK, UK.

Giannis Varelas, Epimenidis Voutsakis, Paraskevi Raftopoulou, Euripides GM Petrakis, and Evangelos E Milios. 2005. Semantic similarity methods in WordNet and their application to information retrieval on the Web. In *Proceedings of the 7th annual ACM international workshop on Web information and data management*, pages 10–16.

Michael J. Wise. 1996. YAP3: improved detection of similarities in computer program and other texts. In *Proceedings of the twenty-seventh SIGCSE technical symposium on Computer science education*, pages 130–134, Philadelphia, Pennsylvania, USA.

Ichiro Yamada, Jong-Hoon Oh, Chikara Hashimoto, Kentaro Torisawa, Jun'ichi Kazama, Stijn De Saeger, and Takuya Kawada. 2011. Extending WordNet with hypernyms and siblings acquired from Wikipedia. In *Proceedings of 5th International Joint Conference on Natural Language Processing*, pages 874–882, Chiang Mai, Thailand.

Torsten Zesch, Christof Müller, and Iryna Gurevych. 2008. Extracting lexical semantic knowledge from Wikipedia and

Wiktionary. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC)*, pages 1646–1652, Morocco.