A Survey of WordNets and their Licenses

Francis Bond
Linguistics and Multilingual Studies
Nanyang Technological University
bond@ieee.org

Kyonghee Paik
Waseda Media Network Center and
Center for Modern Languages, NTU
paikbond@gmail.com

Abstract

This paper surveys currently available wordnets. We measure the effect that license choice has on their usage, measured by the number of citations. Finally, we discuss methods to make wordnets more generally accessible, starting with a shared online server for freely distributable wordnets.

1 Introduction

In this paper we answer two questions: (i) what effect does license choice have on wordnet uptake? and (ii) How can we make wordnets more widely useful? To answer these questions we start off by surveying the current available wordnets and end up trying to make a multilingual wordnet server.

This paper was mainly inspired by two things. The first was a discussion with researchers in Europe who were presenting their work on linking semantic classes between French and Italian. They were using semantic classes derived from clustering Europarl and we asked why they didn't use WordNet. Their answer was that it wasn't available for their languages. Our first reaction was shock — of course there are wordnets for French and Italian. On relection, we realized that in fact, they are somewhat hard to find. If you search for downloadable wordnets for these two languages, you end up at the ELRA page, where you are charged a considerable amount of money for out-of-date versions from the EuroWordNet project. If you know what to look for, you can find free (at least for research) up-to-date versions of wordnets for French and Italian, called WOLF and MultiWordNet respectively, but they are not obvious.

The second inspiration was discussions within projects we have been involved with building wordnets, where we were trying to decide which license should be used. We found there was very little in the way of qualitative evidence that one license was better than another, and decided to try to produce some ourselves. There are over 40 projects to build wordnets for various languages. They range from the Princeton WordNet of English (Fellbaum, 1998) the original wordnet project which has over 150,000 concepts, to research projects such as those on Bantu or Norwegian, which have yet to release any results. Because there are so many wordnets all sharing a similar structure, but with a wide variety of licenses, they provide good data to look at license use.

Language resources, to be useful, must be both **accessible** (legally OK to use) and **usable** (of sufficient quality, size and with a documented interface) (Ishida, 2006). We address both of these concerns in this paper.

This paper is structured as follows. First, we survey currently available wordnets (§ 2). Based on the results we look at the effect that license choice has on wordnet citations (§ 3). Then we look at building a combined multilingual wordnet based on the available free wordnets (§ 4). We finish with some discussion (§ 5) and then conclude.

2 A Survey of WordNets and their Licenses

We have compiled a survey of wordnet projects, and found the license for the projects that have released data (see Table 1). This table shows the project name, languages included, number of synsets, first release, license, canonical citation and number of citations for the canonical citation.

Roughly a third of the wordnets are **open source**, that is, free and redistributable with

no constraints. The most common license is a variant of original Princeton Wordnet License (a modified MIT license), the rest are free licenses such as the LGPL, GNU-FDL, GPL or CC BY. These are all free, open-source licenses according to widely accepted definitions such as the Open Source Definition,[1] Debian Free Software Guidelines[2] or the GNU project.[3]

Roughly a third are available **free for research**, but they cannot be redistributed or have some restriction on their use and are thus not open source licenses. Most of these also offer a separate license for commercial use.

Finally, there are roughly a third of the wordnets which are **non-free**, costing money even for research use, although generally with a reduced cost. Some wordnets have both free and non-free versions. In particular, wordnets produced through Euro WordNet are all sold by the European Language Resources Association (ELRA) even if exactly the same data or updated newer versions are also available for free.

There is a trend for newer projects to use open licenses (such as Japanese, Finnish and Thai) and for older projects to re-release their data under more open licenses (such as German and Catalan).

To compile this data we took as our starting point the table of WordNet projects *Wordnets in the World*[4] maintained by the Global WordNet Association (GWA). In addition to projects listed here, we have added other projects that we discovered through mailing lists or conferences.

As our interest is mainly in NLP applications, we have not included wordnets which have an online interface but no information on how to obtain the whole database (such as Latin, Nepali, Portuguese, and many others). We have done our best to survey as many wordnets as we can, but apologize in advance if we have missed any. There are some projects (e.g., Albanian) where we could actually download the data, but could not find information on the license, we have omitted them from the table.

[1] `www.opensource.org/docs/definition.php`

[2] `www.debian.org/social_contract.html#guidelines`

[3] `www.gnu.org/licenses/license-list.html`

[4] `www.globalwordnet.org/gwa/wordnet_table.htm`

Size	Date	Open	Free	Non free
Large	2008	Japanese	Dutch	
		24	19	
Large	2009	Danish/Thai		Korean
		8/4		5
Small	2008	French	Slovenian	Bulgarian
		22	13	3

Table 2: Similar wordnets with different licenses

We sent out a questionnaire to get more data for the paper. Entries in the table marked with a * are based on the questionnaire responses. The questionnaire is given in Appendix A.

We have created a map based on this data (Figure 1). The map is based on countries, which means that language/mapping involves some subjective judgment.[5] Languages with an open source wordnet are shown in green, with a free for research wordnet in blue, and a non-free wordnet in brown. The higher the number of synsets, the lighter the color. When there are two wordnets with different licenses, we have used the most complete version as representative for that languages.

The map shows that much of the world has at least some wordnet for it, although the coverage of Africa and central Asia is still very incomplete.

3 Assessing the Effect of License Choice on WordNet Citations

In order to compare effect of license choice on number of citations, we compare a few similar size projects released at similar times in Table 2. Any comparisons done in this way are inherently noisy — Google Scholar counts may be wrong and citations may depend on any number of factors, including the novelty of the construction method, the number of people working on the language, the accessibility of the publication venue and so forth. However, comparing the most similar pairs we can, in general the more free version has the most citations.

There are some exceptions, the Chinese wordnet (Xu et al., 2008) has no citations, despite it being large and **free for research**. We suspect that this is because you must email

[5] For example, we have marked Spain as having free coverage thanks to the Catalan Wordnet.

Name	Language	# Synsets	Release	License	Citation	Count
Open Source						
Princeton WN*€	English	155,000	1991	WordNet	Fellbaum (1998)	6,821
FinnWordNet	Finnish	117,700	2010	WordNet	Lindén and Carlson. (2010)	0
Russian WN	Russian	117,000	2004	Wordnet	Balkova et al. (2008)	15
Thai Wordnet	Thai	73,593	2007	WordNet	Thoongsup et al. (2009)	4
DanNet*	Danish	65,000	2008	WordNet	Pedersen et al. (2009)	8
Japanese WN*	Japanese	57,000	2009	WordNet	Isahara et al. (2008)	24
Catalan WN*	Catalan	42,000	1999	GPL	Benítez et al. (1998)	17
LSG	Irish Gaelic	32,742	?	GNU FDL	`http://borel.slu.edu/lsg/`	—
Hindi WN	Hindi	28,687	?	GNU FDL	Jha et al. (2001)	10
WOLF	French	22,000	2009	Cecill-C†	Sagot and Fišer (2008)	22
Wordnet Bahasa*	Malay, Indonesian	20,000	2011	MIT	Nurril Hirfana et al. (2011)	—
Spanish WN*⊙€	Spanish	15,556	2006	LGPL	Farreres et al. (1998)	65
Catalan WN*⊙€	Catalan	15,556	2006	LGPL	Benítez et al. (1998)	17
Arabic WN*	Arabic	11,269	2008	CC BY SA	Black et al. (2006)	28
Hebrew WN*	Hebrew	5000	2006	GPL	Ordan and Wintner (2007)	0
Free for Research						
Chinese WN*	Chinese	115,424	2008	res/com	Xu et al. (2008)	0
KorLex*⊙	Korean	90,000	2007	res/com	Yoon et al. (2009) (nouns)	—
Spanish WN*€	Spanish	62,000	1999	res/com	Farreres et al. (1998)	65
Cornetto*€	Dutch	70,371	2009	res/com	Vossen et al. (2008)	19
GermaNet*€	German	69,594	2011	res/com	Kunze and Lemnitzer (2002)	52
MultiWN*€	Italian	38,877	2008	res/com	Pianta et al. (2002)	143
MWN*	Macedonian	33,276	2010	CC BY NC	Saveski and Trajkovski (2010)	0
Ro-WordNet*	Romanian	30,000	soon	no-deriv.	Tufiş et al. (2008)	9
Czech WN *€	Czech	29,000	1999	res/com	Pala and Smrž (2004)	34
SloWnet*	Slovene	20,000	2010	CC BY NC SA	Fišer and Sagot (2008)	13
Non Free (Available for Research)						
KorLex*	Korean	130,878	2007	res/com	Yoon et al. (2009)	5
Estonian*€	Estonian	47,000	—	ELRA	Kerner et al. (2010)	0
EuroWordNet					Vossen (1998)	728
Dutch	Dutch	44015	1999	ELRA	ELRA-M0016	
Spanish	Spanish	23370	1999	ELRA	ELRA-M0017	
Italian	Italian	48529	1999	ELRA	ELRA-M0018	
German	German	15,132	1999	ELRA	ELRA-M0019	
French	French	22,745	1999	ELRA	ELRA-M0020	
Czech	Czech	22,745	1999	ELRA	ELRA-M0021	
Estonian	Estonian	9,317	1999	ELRA	ELRA-M0022	
ItalWordNet	Italian	49,360	1999	ELRA	ELRA-M0018	
BasqWN	Basque	30,281	?	ELRA	Pociello et al. (2011)	0
BulNet*⊙	Bulgarian	23,715	2004	ELRA	ELRA-M0041 (Koeva, 2008)	3

Table 1: Catalog of WordNets

*Results from our survey
⊙A subset released under a less restrictive license
€A version from EuroWordNet is also available from ELRA
† A variant of the LGPL
res/com means that it is available under different license for research and commercial use
Release is the first release under this license

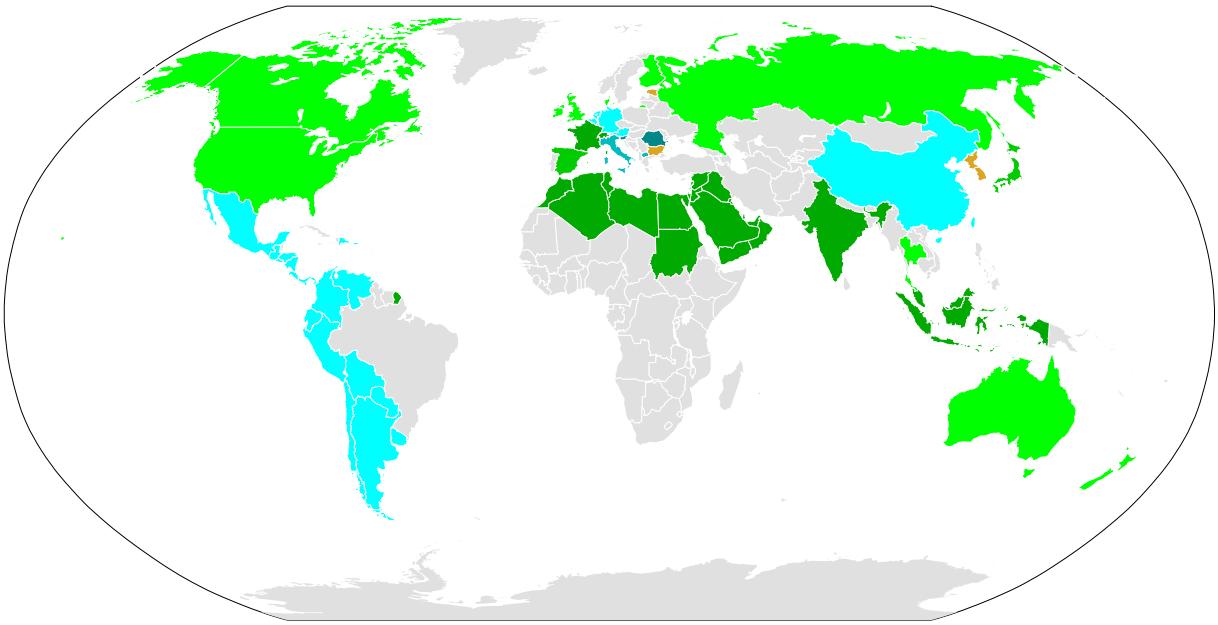


Figure 1: Map of Countries showing WordNet availability

Countries with **open source wordnets** are in green; **free for research** wordnets are in blue; **non free** wordnets are in brown. The lighter the color, the more synsets.
Citation counts from Google Scholar (accessed on 2011-09-23)

and ask for permission to use it, which is a substantial barrier to use. The Italian wordnet (Pianta et al., 2002) has a very high number of citations. In this case it was developed as part of a multilingual wordnet with several other languages, thus giving it a large citation group. Finally, the Thai wordnet (Thoongsup et al., 2009), has relatively few citations, in this case it is also a part of a large project (the Asian Wordnet: Sornlertlamvanich et al. (2008)) which gets more citations (10).

Even with all these caveats, we think that the data supports the unsurprising result that the more open the license a wordnet is released under, the more likely it is to be used (or at least cited). In other word, uptake of a resource depends on how **usable** (legally accesible) a resource is.

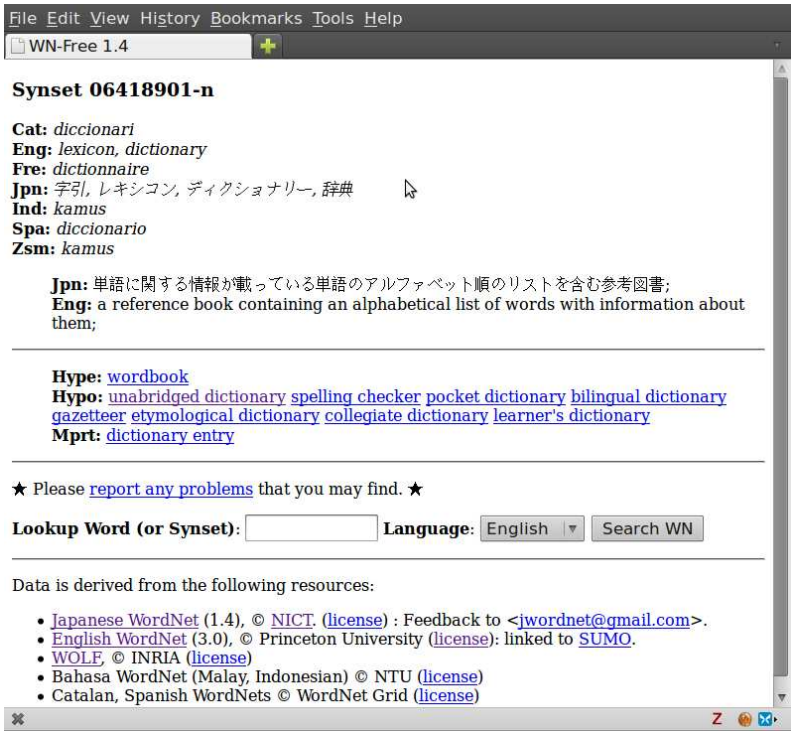
4 Construction of an Open Multi-lingual WordNet

In order to make the wordnets more **accessible**, we have started to build a simple SQL server with information from those wordnets whose licenses allows us to do so. We show a screen shot in Figure 2. Most of these wordnets are based on the **extend** model, basically adding lemmas in new languages to existing English synsets. Therefore, adding a new language is a case of just adding new lemmas to the synsets (annotated with their language). In theory, this should be easy.

In practice, adding a new language turned out to be difficult for two reasons. The first problem was that none of the wordnets we surveyed updated their structure when the English wordnet did. In order to combine them into a single multilingual structure, we had to map to a common version. The second problem was the incredible variety of formats that the wordnets are distributed in. Almost every project used a different format and thus required a new script to convert it. These two problems mean that, even if a wordnet is legally available, there is still a technical hurdle before it becomes easily accessible.

The first problem can largely be overcome using the mapping scripts from Daude et al. (2003). Mapping introduces some distortions, in particular, when a synset is split, we chose to only map the translations to the most probable mapping, so some new synsets will have no translations.

The second problem we are currently solving



Catalan, English, French, Japanese, Indonesian, Spanish and Malaysian wordnets stored in a common database searchable database.

Figure 2: Screenshot of Free Wordnet Lookup

through brute force, writing a new script for almost every new wordnet we add. We discuss better possible solutions in Section 5.

The server currently includes English (Fellbaum, 1998), Japanese (Bond et al., 2008); French (Sagot and Fišer, 2008); Indonesian and Malaysian (Nurril Hirfana et al. 2010); Catalan and Spanish from the Global WordNet Grid (Fig 2.) The wordnets are all in a shared `sqlite` database with a `PERL` cgi server using the wordnet module produced by the Japanese WordNet project (Isahara et al., 2008).

We plan to add further free wordnets. Additional wordnets with licenses that allow us to serve the data exist for Arabic, Danish, Finnish, Gaelic, Hebrew, Hindi, Macedonian, Slovenian and Thai (see Table 1 or 3).

We also have a script that outputs wordnets from our database into either tab separated values, where they can be used by the Natural Language Tool Kit[6] (Bird et al., 2009) or the emerging standard of WordNet-LMF (Lexical Markup Framework: Soria et al., 2009). Finally, we will also make the SQL database available. Licenses that allow redistribution of derivative works allow people to make the entire lexicons available in any format, thus greatly improving their usefulness.

Notes on the various formats and conversions

Our conversion scripts basically reduce each wordnet to a list of synset-lemma pairs, which we then map to the English 3.0 synsets. We currently lose any extra information about, for example, other morphological forms of words (we expect the morphological analyzers to give us lemmas). We also lose any synsets not in the English 3.0 wordnet.

T/CSV (Tab/Comma Separated Values) The Japanese Wordnet is already available as a table of synset-lemma pairs, and a third field that gives a confidence score, so we can use it immediately. The Bahasa Wordnet has an additional field that says whether each synset-lemma pair can be used in Malay or Indonesian or either (98% of entries are ac-

[6]With the extensions that were added with the Japanese translation (Bird et al., 2010).

Wordnet	Ver	Format
Arabic	?	pwn
Bahasa✓	3.0	tsv
Catalan✓	1.6	gwa-xml
Danish	2.0	csv, owl
English	3.0	pwn*
Finnish	3.0	text tables
French (WOLF)✓	2.0	xml
Irish		
Japanese✓	3.0	tsv
Hebrew		xml
Hindi	?	pwn
Spanish✓	1.6	gwa-xml
Thai	3.0	LMF

Table 3: Free WordNets and their formats

*Read using the perl WordNet::QueryData module (Rennie, 2000)

✓ Conversion script built

Ver. is the corresponding English version

ceptable in either language). We pre-process the wordnet to output two wordnets, one for Malay and one for Indonesian. The Danish wordnet is similar to the Japanese wordnet.

GWA-XML A subset of the Spanish and Catalan wordnets are released under the LGPL. They are released in pseudo XML. We convert them to actual XML by fixing the encoding of the quotation marks and adding a tag surrounding the whole file (`<wn>...</wn>`[7]). French (WOLF) and Hebrew are released in a very similar format. This format is based on the de facto standard established by the Euro WordNet project.

We have a simple script that pulls out the synset-lemma pairs from this XML (and ignores everything else).

LMF To get a mapping from a lemma to the English synset in LMF theoretically requires two mappings (lemma to language synset, language synset to English synset). In practice, for the current version of the Thai Wordnet, the Thai synset ID is always the same as the English 3.0 ID. We take advantage of this to extract the information we want with a very simple script.

[7]We have reported these problems upstream.

The remainder of the wordnets all use quite different formats. We are hoping to persuade each individual project to also output the data in a more universal format.

5 Discussion and Further Work

In general, the freer the license, the more a wordnet appears to be used. Therefore, for projects whose goal is to produce a resource that will be widely used, the freer the license the better.

Access to multiple wordnets would, of course, be simplest if everyone used the same format. This would also make tool sharing easier and perhaps reduce some of the current duplicated effort. Currently there seem to be two front runners: Wordnet-LMF (Soria et al., 2009), which is explicitly designed as an interchange format, and Wordnet SQL.[8] Both of them are easily processed with existing interfaces (for XML and SQL) unlike the PWN format, which is very specialized to wordnet. However, different projects already have invested effort into their current interfaces and formats, so may not change quickly. In addition, it is often easier to get funding to build resources that to maintain them.

Given this, a more realistic medium term goal for increasing access to multiple wordnets is to encourage conversion from whatever local representation to a shared representation, such as Wordnet-LMF. Because accurate conversion relies on knowledge of each wordnet's individual structure, it would be safer if each project did this conversion themselves. As a start, we will release our conversion scripts.

In the near future, we also plan to improve our conversion scripts so as to also add new synsets when they exist, although there is currently no way to link them across different language pairs. This problem was solved with the Inter-Lingual Index in EuroWordNet (Vossen, 1998), but currently there is not yet a single index shared by all projects.

All of the data in our catalog, and the map, is being fed back to the *Wordnets in the World* page. We have already been sending additions and corrections to the maintainers throughout the project. In particular, we found the contact details were out of date for 5 out of 45

[8]`wnsql.sourceforge.net/uml2.html`

projects, and are in the process of finding the current contacts for these projects. We are also adding projects that we know of through the Asian WordNet meetings to the list.

Finally, on a positive note, regardless of the actual license, researchers are generally very willing to share their data, and will often make it available on request, or even link to it online, even though the license does not, strictly speaking allow this. While this is very welcome, receiving data without a proper license does not legally allow its use, and thus does not lead to more reproducible research.

6 Conclusions

We have surveyed the current coverage of wordnets, both in terms of size and license. Many of the world's most widely used languages now have wordnets, although not all of them are freely available, and lack of standard interfaces and data formats makes them hard to access. We have made a first step to increasing accessibility by converting free wordnets to a common format. We show that, in general, wordnets released with freer licenses are cited more often.

Acknowledgments

We would like to thank the compilers of the Global WordNet Association's *Wordnets in the World* page as well as everyone who responded to the survey. This research was funded by the Creative Commons Catalyst Grant: *Assessing the effect of license choice on the use of lexical resources.*

References

Valentina Balkova, Andrey Sukhonogov, and Sergey Yablonsky. 2008. Russian wordnet: From UML-notation to internet/intranet database implementatiom. In Petr Sojka, Karel Pala, Pavel Smrž, Christiane Fellbaum, and Piek Vossen, editors, *Proceedings of the Second International WordNet Conference — GWC 2004*, pages 31–38. Brno.

L. Benítez, S. Cervell, G. Escudero, M. López, G. Rigau, and M. Taulé. 1998. Methods and tools for building the Catalan wordnet. In *ELRA Workshop on Language Resources for European Minority Languages*. Granada, Spain.

Stephen Bird, Ewan Klein, and Edward Loper. 2009. *Natural Language Processing with Python*. O'Reilly. (`www.nltk.org/book`).

Stephen Bird, Ewan Klein, and Edward Loper. 2010. *Nyumon Shizen Gengo Shori [Introduction to Natural Language Processing]*. O'Reilly. (translated by Hagiwara, Nakamura and Mizuno).

W. Black, S. Elkateb, H. Rodriguez, M. Alkhalifa, P. Vossen, A. Pease, M. Bertran, and C. Fellbaum. 2006. The Arabic wordnet project. In *Proceedings of LREC 2006*.

Francis Bond, Hitoshi Isahara, Kyoko Kanzaki, and Kiyotaka Uchimoto. 2008. Boot-strapping a WordNet using multiple existing WordNets. In *Sixth International conference on Language Resources and Evaluation (LREC 2008)*. Marrakech.

Jordi Daude, Lluis Padro, and German Rigau. 2003. Validation and tuning of Wordnet mapping techniques. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP'03)*. Borovets, Bulgaria.

X. Farreres, G. Rigau, and H. Rodríguez. 1998. Using wordnet for building wordnets. In *COLING-ACL Workshop on Usage of WordNet in Natural Language Processing Systems*. Montreal, Canada.

Christine Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.

Darja Fišer and Benoît Sagot. 2008. Combining multiple resources to build reliable wordnets. *Text, Speech and Dialogue*, LNCS 2546:61–68.

Hitoshi Isahara, Francis Bond, Kiyotaka Uchimoto, Masao Utiyama, and Kyoko Kanzaki. 2008. Development of the Japanese WordNet. In *Sixth International conference on Language Resources and Evaluation (LREC 2008)*. Marrakech.

Toru Ishida. 2006. Language grid: An infrastructure for intercultural collaboration. In *IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06)*, pages 96–100. URL `http://langrid.nict.go.jp/file/langrid20060211.pdf`, (keynote address).

S. Jha, D. Narayan, P. Pande, and P. Bhattacharyya. 2001. A wordnet for Hindi. In *International Workshop on Lexical Resources in Natural Language Processing*. Hyderabad.

Kadri Kerner, Heili Orav, and Sirli Parm. 2010. Growth and revision of Estonian wordnet. In Pushpak Bhattacharyya, Christiane Fellbaum, and Piek Vossen, editors, *5th Global Wordnet Conference: GWC-2010*, pages 198–202. Mumbai.

Sv Koeva. 2008. Derivational and morphosemantic relations in Bulgarian wordnet. In *Intelligent Information Systems, XVI*, pages 359–389. Academic Publishing House, Warsaw.

C. Kunze and L. Lemnitzer. 2002. Germanet — representation, visualization, application. In *LREC*, pages 1485–1491.

Krister Lindén and Lauri Carlson. 2010. Finnwordnet — wordnet påfinska via översättning. *LexicoNordica — Nordic Journal of Lexicography*, 17:119–140. In Swedish with an English abstract.

Nurril Hirfana Mohamed Noor, Suerya Sapuan, and Francis Bond. 2011. Creating the open Wordnet bahasa. In *Proceedings of the 25th Pacific Asia Conference on Language, Information and Computation (PACLIC 25)*. Singapore. (to appear).

Noam Ordan and Shuly Wintner. 2007. Hebrew wordnet: a test case of aligning lexical databases across languages. *International Journal of Translation*, 19(1):39–58.

Karel Pala and Pavel Smrž. 2004. Building Czech wordnet. *Romanian Journal of Information Science*, 7:79–88.

B.S Pedersen, S. Nimb, J. Asmussen, N. Sørensen, L. Trap-Jensen, and H. Lorentzen. 2009. DanNet — the challenge of compiling a wordnet for Danish by reusing a monolingual dictionary. *Language Resources and Evaluation*.

Emanuele Pianta, Luisa Bentivogli, and Christian Girardi. 2002. Multiwordnet: Developing an aligned multilingual database. In *In Proceedings of the First International Conference on Global WordNet*, pages 293–302. Mysore, India.

Elisabete Pociello, Eneko Agirre, and Izaskun Aldezabal. 2011. Methodology and construction of the Basque wordnet. *Language Resources and Evaluation*, 45(2):121–142.

Jason Rennie. 2000. Wordnet::querydata: a Perl module for accessing the WordNet database. http://www.ai.mit.edu/people/jrennie/WordNet.

Benoît Sagot and Darja Fišer. 2008. Building a free French wordnet from multilingual resources. In European Language Resources Association (ELRA), editor, *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*. Marrakech, Morocco.

M. Saveski and I Trajkovski. 2010. Automatic construction of wordnets by using machine translation and language modeling. In Jerneja Zganec Gros. Ljubljana Tomaz Erjavec, editor, *In Proceedings of Seventh Language Technologies Conference, 13th International Multiconference Information Society*, volume C.

Claudia Soria, Monica Monachini, and Piek Vossen. 2009. Wordnet-LMF: fleshing out a standardized format for wordnet interoperability. In *Second International Workshop on Intercultural Collaboration (IWIC-2009)*. Stanford.

Virach Sornlertlamvanich, Thatsanee Charoenporn, Kergrit Robkop, and Hitoshi Isahara. 2008. KUI: Self-organizing multi-lingual wordnet construction tool. In Attila Tanács, Dóra Csendes, Veronika Vincze, Christiane Fellbaum, and Piek Vossen, editors, *4th Global Wordnet Conference: GWC-2008*, pages 417–427. Szeged, Hungary.

Sareewan Thoongsup, Thatsanee Charoenporn, Kergrit Robkop, Tan Sinthurahat, Chumpol Mokarat, Virach Sornlertlamvanich, and Hitoshi Isahara. 2009. Thai wordnet construction. In *Proceedings of The 7th Workshop on Asian Language Resources (ALR7), Joint conference of the 47th Annual Meeting of the Association for Computational Linguistics (ACL) and the 4th International Joint Conference on Natural Language Processing (IJCNLP),*. Suntec, Singapore.

Dan Tufiş, Radu Ion, Luigi Bozianu, Alexandru Ceauşu, and Dan Ştefănescu. 2008. Romanian wordnet: Current state, new applications and prospects. In *Proceedings of the 4th Global WordNet Association Conference*, pages 441–452. Szeged.

P. Vossen, I. Maks, R. Segers, and H Van der Vliet. 2008. Integrating lexical units, synsets and ontology in the Cornetto database. In *LREC 2008*. European Language Resources Association (ELRA), Marrakech, Morocco.

Piek Vossen, editor. 1998. *Euro WordNet*. Kluwer.

Renjie Xu, Zhiqiang Gao, Yuzhong Qu, and Zhisheng Huang. 2008. An integrated approach for automatic construction of bilingual Chinese-English WordNet. In *3rd Asian Semantic Web Conference (ASWC 2008)*, pages 302–341.

Aesun Yoon, Soonhee Hwang, Eunroung Lee, and Hyuk-Chul Kwon. 2009. Construction of Korean wordnet KorLex 1.5. *Journal of KIISE: Software and Applications*, 36(1):92–108.

A Questionnaire about WordNet licensing

This is the questionnaire we sent out, reformatted to fit in one column. We also added a filled-out sample.

```
Please send to: xxx@yyy

More information can be found at:
<http://zzz.yyy>

----------------------------------------
Questions:

* Resource Name:

* Language(s) Described:

* Developer(s):
(Institution)

* Contact Person(s):
(name and email please)

* URL:

* Date of (planned) release:

* License:

* Reason for choosing license:

* Current Size (in synsets):

* Canonical citation(s):

* Funding Source(s):

* Main users:

----------------------------------------
Please answer the questions and send the
questionnaire to <xxx@yyy>.  If you have
any other questions or comments, feel free
 to ask us.

If you would like us to not add the data
to the GWA page, please let us know.
Otherwise we will do so.

We estimate that filling in this form
should take around 15-20 minutes.
----------------------------------------
```