

Knowledge Sources for Bridging Resolution in Multi-Party Dialog

Mark-Christoph Müller¹, Margot Mieskes², Michael Strube³

¹Ubiquitous Knowledge Processing Lab, Technische Universität Darmstadt, Germany
<http://www.ukp.tu-darmstadt.de/people/chmark/>

²European Media Laboratory GmbH, Heidelberg, Germany
<http://www.eml-d.de/english/homes/mieskes/>

³EML Research gGmbH, Heidelberg, Germany
<http://www.eml-research.de/~strube>

Abstract

In this paper we investigate the coverage of the two knowledge sources WordNet and Wikipedia for the task of bridging resolution. We report on an annotation experiment which yielded pairs of bridging anaphors and their antecedents in spoken multi-party dialog. Manual inspection of the two knowledge sources showed that, with some interesting exceptions, Wikipedia is superior to WordNet when it comes to the coverage of information necessary to resolve the bridging anaphors in our data set. We further describe a simple procedure for the automatic extraction of the required knowledge from Wikipedia by means of an API, and discuss some of the implications of the procedure's performance.

1. Introduction

It is commonly accepted that the resolution of bridging anaphora requires access not only to *linguistic*, but also to *common-sense* or *world* knowledge (Clark, 1975).¹ In automatic approaches to bridging resolution, this knowledge is typically provided either in the form of a custom-built and thus domain-dependent knowledge source (e.g. by Hahn et al. (1996)), or in the form of a general-purpose lexical database like WordNet (e.g. by Poesio et al. (1997)). In the latter case, the limited coverage of the respective knowledge source is often problematic. In this paper, we investigate the applicability of Wikipedia as an alternative knowledge source for bridging resolution. Another novelty is that we work on a corpus of unrestricted *multi-party* dialog, while earlier work on bridging in dialog (Nissim et al., 2004) has only considered two-party dialog.

2. Bridging Resolution

Bridging is commonly regarded as a type of anaphoric relation similar to e.g. coreference, on the grounds that both coreference and bridging anaphors refer to some hearer-old entity (Prince, 1992) in a text or dialog. The difference is that while an anaphor and its *coreferent* antecedent are linked by virtue of referring to the same referent, the referential or semantic relation between an anaphor and its *bridging* antecedent can be one of a much larger set.² Computational bridging resolution has not been extensively attempted so far. Most of the few existing systems (cf. above) fall short of being applicable to unrestricted input.

¹The work reported in this paper was done while the first and the second author were affiliated with EML Research gGmbH.

²One also finds a slightly different definition of bridging which includes cases where antecedent and anaphor are actually coreferent, but where the anaphor does not contain the same lexical head as the antecedent (e.g. *the house* - *the building*). For reasons of clarity, we do not follow this definition, but restrict the term *bridging* to non-coreferential cases.

3. Annotations

3.1. Experiment

We performed an annotation experiment with two naive annotators³ on ten randomly selected ICSI Meeting Corpus transcripts (Janin et al., 2003). The ICSI Meeting Corpus is a collection of 75 manually transcribed English-language group discussions of about one hour each. The number of participants in each discussion ranges from three to ten speakers, averaging six. Participants include male and female speakers. There is also a considerable number of non-native speakers of English. The discussions are real, unstaged meetings on various, quite technical topics. Most of the discussions constitute regular weekly meetings, and while each meeting normally centers around a few topics only, the meetings are in principle unconstrained. For bridging resolution, the ICSI Meeting Corpus thus constitutes a rather challenging data basis.

Our annotation manual was based on the hierarchical scheme used by Nissim et al. (2004). This scheme distinguishes the three top-level categories *old*, *mediated*, and *new* (Strube & Hahn, 1999), and optional sub-categories for *old* and *mediated*. For our experiments, we considerably simplified the annotation scheme in several respects. One simplification was that our annotation scheme completely ignored the sub-categories for *mediated* and *old*, mainly because the definition and operationalization of these categories for our naive annotators turned out to be difficult. For example, the sub-category *mediated-situation* was defined by Nissim et al. (2004) on the basis of FrameNet and WordNet lookup, which turned out to be impractical for our annotators. Another difference was that in our annotation, the pronoun *it* and the demonstratives *this* and *that* were to be ignored by

³One female computational linguistics undergrad student, one female psychology grad student. Both annotators were non-native speakers of English.

the annotators, because they had already been studied in a previous annotation experiment (Müller, 2007). We also ignored all first- and second-person pronouns. Among other things, this allowed us to disregard the problem of *generic* pronoun identification (Gupta et al., 2007). In effect, ignoring all the mentioned pronouns led to a major decrease in instances of category *old*. We argue that it is inappropriate to include pronouns (like e.g. Nissim et al. (2004) do), because they are trivial to annotate and apparently can boost reliability scores without really improving the reliability on the non-trivial cases. Generally, the focus of our annotation experiments was more on eliciting intuitive knowledge about inferrable relations from naive annotators, rather than on the precise specification of these relations. The annotation was performed with the annotation tool MMAX2 (Müller & Strube, 2006).

The category *old* was to be applied to noun phrases whose referents had already been introduced into the discourse, and who were therefore coreferential with some preceding antecedent. For *old* noun phrases, the annotators were also asked to identify this antecedent, and to link it to the *old* noun phrase (i.e. the anaphor) by means of co-indexing. In the following example, *these systems* was annotated as *old* by both annotators, and *content management systems* was identified as its antecedent.⁴

MN059: OK, so, um, what I started looking at, uh, to begin with is just uh, [content management systems]_i uh, i- i- in general.

(Some intervening utterances by the same speaker)

MN059: Now, if you sort of put on your semantic glasses, uh you say, well that's not all that easy, because there's an implicit um, uh, assumption behind that is that uh, all the users of this system share the same interpretation of the keyword and the same interpretation of uh, whichever taxonomy is used, and uh, I think that's a - that's a very - that's a key point of [these systems]_i and they sort of always brush over this real quickly without really elaborating much of that and uh - (Bed017)

The category *mediated*, on the other hand, was to be applied to noun phrases whose referents had not yet been introduced, but which were inferrable from (1) previously mentioned referents or (2) general knowledge. For the annotation of the first type of *mediated* noun phrases, the annotators had at their disposal a functionality of the MMAX2 tool which allowed to have the noun phrase point to the bridging antecedent, i.e. to the closest noun phrase mention (if any) of the referent sponsoring the bridging relation. Consider the following example, in which both annotators classified *the data* as *old* and *the data collection* as its bridging antecedent.

MN015: Um, *outbreath* uh in a - in a smaller group we had uh, talked and decided about continuation of [the data collection].

FN050: *mike noise*

MN015: So Fey's time with us is almost officially over, and she brought us some thirty subjects and, t- collected [the data]_{old}, and ten dialogues have been transcribed and can be looked at. (Bed017)

For the second type of *mediated* noun phrases, no textual bridging antecedent is available by definition. Rather, bridging anaphors of this type are mediated either by the physical dialog context (e.g. *this room*, *the print outs* (that have been handed to the participants)), or by general knowledge shared by all dialog participants (e.g. *the data collection* or *the speech community*).

Finally, the category *new* was applied to noun phrases that were neither *old* nor *mediated*, **if** the noun phrase also served as the antecedent of at least one *old* or *mediated* noun phrase. This constraint was used in order to reduce the number of *new* markables by ignoring e.g. the large number of singletons, i.e. noun phrases that are mentioned only once. As a result, the total number of *new* noun phrases is much lower than e.g. in Nissim et al. (2004). Also, since identification of *new* is closely coupled with the much more difficult identification of coreferential relations, the inter-annotator agreement regarding the identification of *new* noun phrases can be expected to be low. Cases where one or both of the annotators skipped a noun phrase (left it unannotated as *default*) were ignored.

3.2. High-Level Agreement

We used the κ statistic (Fleiss, 1971) to calculate the inter-annotator agreement for the three-fold classification in *old*, *mediated*, and *new* in each of the ten dialogs. The result can be found in Table 1.

	Old	Mediated	New	all
Bed016	.78	.71	-.02	.71
Bed017	.77	.59	.51	.66
Bmr001	.80	.59	.16	.63
Bmr002	.78	.69	.40	.71
Bns003	.73	.55	.16	.59
Bro003	.68	.57	.08	.60
Bro004	.77	.54	.19	.60
Bro005	.79	.69	.29	.71
Bsr001	.76	.69	.49	.71
Btr001	.79	.73	.14	.74

Table 1: Agreement (κ) for three-fold classification.

It can be seen that the class *old* can be assigned with good reliability, while the class *mediated* is above the common .67 threshold in only half of the cases. The reliability for the class *new*, finally, is very low, as was to be expected from the way the class was defined.

The confusion matrix in Table 2 is a way to inspect the quantitative aspect of the agreement and disagreement of the two annotators.

The bold figures on the diagonal are the cases of agreement, and the cardinality of the figures reflect the degree of agree-

⁴Here and in the following examples, other annotations are left out for clarity.

	Old	Mediated	New	Anno 1
Old	2552	221	10	2783
Mediated	137	743	13	893
New	23	163	44	230
Anno 2	2712	1127	67	3906

Table 2: Confusion matrix.

ment. The two annotators agree in 2552 of the 2943 cases (i.e. in 86.71%) in which at least one of them assigned the class *old*. For *mediated*, the agreement is still 58.12%, while for *new*, it is merely 17.39%.

3.3. Agreement of Bridging Antecedents

As stated in the introduction, the focus of this paper is on an empirical evaluation of WordNet and Wikipedia as knowledge sources for the resolution of bridging expressions. More precisely, we want to investigate whether they contain information that allows to establish a bridging relation between one expression as a potential bridging expression and another expression as its potential bridging antecedent. In doing so, we deliberately restricted ourselves to definite noun phrases as potential bridging expressions, i.e. those beginning with *the*, *this*, *that*, *these*, or *those*. This was motivated by the fact that indefinite noun phrases (i.e. those with no article or with *a/an*) are normally not anaphoric, but referentially self-contained. During coreference resolution, indefinite noun phrases are only considered as potential antecedents, but not as potential anaphors that require resolution. Definite noun phrases, in contrast, are by default taken to be potentially anaphoric, and normally trigger a coreference resolution attempt. Definite noun phrases that actually are non-anaphoric bridging expressions are a source of error in coreference resolution because they can give rise to incorrect coreference relations.

The raw data produced by the manual annotation experiment was processed as follows. First, we extracted from the ten dialogs all 743 instances which both annotators annotated as *mediated*. Of these, 324 were definite noun phrases. For each of the definite noun phrases annotated by each of the two annotators, we then extracted from the manual annotations the bridging antecedent identified by the respective annotator (if any). For the comparison of these bridging antecedents, two degrees of identity were defined: *Same antecedent* means that both annotators identified the same noun phrase token as the bridging antecedent for a given noun phrase, while *same head* means that they identified different tokens, but that these tokens contained the same lexical head. We argue that the inclusion of the second degree of identity makes sense for the following reasons: First, different occurrences of the same noun phrase may be coreferential, in which case there is no difference between *same antecedent* and *same head*. Second, in order to establish for a given definite noun phrase that it is a bridging expression, it is sufficient to identify *any* noun phrase capable of sponsoring a bridging relation to it. In other words: What is important here is that the annotators agreed on the semantic nature of the antecedent, not necessarily on the actual token.

Table 3 shows the distribution of types.

	All	Definites only
Same antecedent	86	70
Same head	22	14
Different antecedents	31	27
One antecedent missing	129	72
Two antecedents missing	475	141
Σ	743	324

Table 3: Types of *Mediated* NPs and their Antecedents.

Among the 324 definite noun phrases classified as *mediated*, we found only 70 for which both annotators identified the same antecedent, and a mere 14 in which they identified different antecedents with the same head. In 27 cases, each annotator identified a different antecedent with a different head for a given noun phrase classified as *mediated*, and in 72 cases, only one of the annotators identified any antecedent at all. For a huge number of cases (141), none of the two annotators identified any antecedent at all. Thus, we were left with only 84 pairs of bridging expressions and antecedents that both annotators agreed upon. As was mentioned in the beginning of this section, our annotators were relatively free in assigning the class *mediated*. As a result, the identified pairs exemplify diverse semantic relations, only some of which require world knowledge for their resolution. The first two columns of Table 4 contain a list of these pairs.

Covered relations in the table include part-of (like in *cafe - the floor* or *table - the column / the line*), but also more generally associative relations, like in *field trip - the logistics*.

4. Evaluation of Knowledge Sources

In this section, we first evaluate the coverage of two knowledge sources for bridging resolution. More precisely, we investigate in Section 4.1. whether WordNet (Fellbaum, 1998)⁵ and the English Wikipedia (<http://en.wikipedia.org>) contain information necessary to recover the bridging relation between the antecedents and bridging expressions listed in Table 4. This initial evaluation is done manually and very informally only, because here our main interest is to establish whether the required information is included at all. In doing so, we disregard the problem of how it might be extracted automatically. This problem will then be addressed in Section 4.2.

4.1. Coverage

We used the web-based query interfaces of WordNet and Wikipedia to search for the base form of the bridging antecedents in Table 4, and inspected the retrieved entries with respect to whether they contained a mention of the respective anaphor. The reason for doing it this way (rather than trying to find a mention of the antecedent in the anaphor's entry) is that the association between bridging anaphor and antecedent is normally not symmetrical. As mentioned at the end of Section 3.3. above, the bridging anaphor is often used to refer to something that stands in a relation like e.g. part-of to the antecedent. When we apply this argumentation to the structure of our knowledge bases, it means that

⁵We used the web interface access to version 3.0

Antecedent	Bridging Anaphor	WordNet		Wikipedia	
		Antecedent Entry Found?	Sufficient?	Antecedent Entry Found?	Sufficient?
microphone	batteries	yes	no	yes	yes
microphone	switch	yes	no	yes	no
university	address	yes	no	yes	no
cafe	floor	yes	yes	yes	no
data collection	data	no	-	yes	yes
Bayes-net	input nodes	no	-	yes	no
field trip	logistics	yes	no	yes	no
neural net	training	yes	no	yes	yes
experiment	result	yes	no	yes	yes
table	column	yes	yes	yes	yes
table	line	yes	no	yes	no
problem	answer	yes	no	yes	yes
(two) people	(the) weaker voice	yes	no	yes	no
France	villages	yes	no	yes	no
utterance	beginning	yes	no	yes	no
question	answer	yes	no	yes	yes

Table 4: Bridging Examples based on World Knowledge.

it is more probable for the *microphone* entry to contain a reference to batteries (because some types of microphones are powered by batteries), than it is for the *battery* entry to contain a reference to microphones.

4.1.1. WordNet

For WordNet, we found that all but two antecedents were covered: Only for *data collection* and *Bayes-net* there are no entries in WordNet. However, only for two antecedents there was information in WordNet which allowed to link the bridging anaphor to the antecedent: For *cafe*, a link to *floor* could be established via *cafe is-a building*, *building* has *room* and *room* has *floor*. For *table*, on the other hand, the link was more explicit, as the gloss definition contained a reference to *column* and *row* (but not the also required *line*).

4.1.2. Wikipedia

For Wikipedia, the results are as follows: For all antecedents in Table 4, there is a Wikipedia article, and for seven of these, the article contains information sufficient for recovering the bridging relation. Cases where Wikipedia also fails include *microphone* - *the switch*, *field trip* - *the logistics*, and *table* - *the line*. Not surprisingly, Wikipedia contains a lot of information on computer-related terms, including *Bayes net* and *neural net*.

4.2. Automatic Extraction

The results of the informal investigation in Section 4.1. above show that our two knowledge sources contain sufficient information for resolving many of the cases in our selection of bridging expressions. This is obviously not sufficient, however, because for automatic bridging resolution an automated extraction process is required. In this section, we explore a way of extracting the required information programmatically by means of an API. In doing so, we restrict ourselves to Wikipedia, which the above informal evaluation showed to be superior to WordNet for the task at hand.

4.2.1. Procedure

We implemented a simple prototype system on the basis of the Java Wikipedia Library (JWPL) described in (Zesch et al., 2008). Among many other things, this API allows access to both the plain text of a Wikipedia page and to the linking structure between pages. The latter functionality includes methods to retrieve all *outlinks* of a page, i.e. those words in the page that constitute links to other Wikipedia pages. In our evaluation, we used two types of matching setups: Matching against the full text of the page, and matching against the outlink words only. The latter setup is motivated by the idea that outlink words are of central importance to the concept described on the source page (Mihalcea & Csomai, 2007). Thus, if the linking between Wikipedia pages is sufficiently dense, these links alone can provide high-precision information.

The extraction process worked as follows: From each pair of bridging anaphor and antecedent in Table 4 for which there was an entry in Wikipedia (cf. Table 5), we first extracted from the antecedent the lemmatized head noun *and any preceding adjectives*. POS tagging and lemmatization was done using the TreeTagger (Schmid, 1997). The extraction of both head and adjectival modifiers was necessary for cases like *neural net* in which an adjectival modifier was integral part of the antecedent. We then used the API to retrieve the pertaining Wikipedia entry (if any). If no entry was found, the lookup was repeated with the lemmatized antecedent head noun only. If this still did not return an entry, automatic extraction failed. If an entry was found, we next extracted the lemmatized head noun of the bridging anaphor, and tried to match it against the full text of the antecedent's Wikipedia page and against that page's outlinks only. The results of this automatic extraction procedure are discussed in the next section.

4.2.2. Results

The results of the automatic extraction procedure can be found in Table 5. The following points are worth noting: The automatic extraction procedure fails to find Wikipedia entries for two expressions which, according to the manual

lookup, should be there. Closer inspection revealed the following reasons, which both highlight two characteristics of Wikipedia, and of community-generated knowledge bases in general.

The API uses a local version of Wikipedia which does not contain the entry for *data collection* yet. This entry, according to the history section of the page http://en.wikipedia.org/wiki/Data_collection, was only created on May 30th, 2007.

In both the online version and the API, the retrieval of the entry for *table* returns a disambiguation page, i.e. a page containing links to the several possible readings (or senses) of *table* that are covered in Wikipedia, including e.g. *Table (furniture)* and *Table (database)*. During the manual inspection, selection of the correct reading *Table (information)* was a rather trivial step. For the API, on the other hand, this is not trivial at all. For many ambiguous expressions (but not for *table*), Wikipedia contains community-generated information about default readings. When these are unavailable, the API has no means of determining the reading that is (probably) correct.

Another interesting result relates to the performance of the matching based on the pages and the page outlinks, respectively. For the five entries retrieved automatically, all the required bridging anaphors could be found in the corresponding pages. In only one case, however, a match was found that was *also* an outlink of that page.

5. Conclusion and Future Work

In this paper, we investigated the applicability of Wikipedia as a knowledge source for the resolution of bridging expressions collected in a corpus of unrestricted multi-party dialog. The results of the initial annotation experiment showed that the agreement on the identification of *mediated* expressions, which are a superclass of bridging anaphors, is above the commonly accepted reliability threshold in some cases only. The agreement of the antecedents of the anaphors that both annotators classified as *mediated* was equally low. This low agreement resulted in a rather small data set of pairs of bridging anaphors and antecedents as the basis of our qualitative evaluation.

The findings of this qualitative evaluation indicate that the degree of detail found in Wikipedia articles makes them good candidates for the task of bridging resolution. WordNet, in contrast, suffers from two limitations: The coverage is limited, which is particularly apparent in the field of computer-related terms, of which our corpus happened to contain quite a few. Also, WordNet focusses more strongly on the hierarchical modelling of lexical relations between words, whereas the definitions (in the form of glosses) are rather brief. Very fundamental relations (like *building - room - floor*) are modelled explicitly, which accounts for the one case where WordNet outperforms Wikipedia in terms of coverage. In most cases, however, the required knowledge is less rigid, and thus more naturally found in the (sometimes rather long) freetext Wikipedia articles.

The evaluation of a prototypical automatic extraction procedure showed that the functionality provided by the Wikipedia API employed is in principle sufficient for automatically extracting the required knowledge from

Wikipedia. One of two observed failures was due to different Wikipedia versions (online vs. local), which is not a problem of automatic access *per se*, but which highlights the problems that might arise from highly dynamic, community-generated content. The other case in which automatic extraction failed was more serious: For genuinely ambiguous expressions, there is no easily automated way of selecting the correct of several matching entries. Even when default readings are encoded in Wikipedia (which is not always the case), these are only heuristics based on the most common underlying word sense.

Future work, therefore, will have to include the investigation of means for automating the processing of Wikipedia disambiguation pages, which showed to be problematic in the automatic extraction. Existing approaches to be evaluated include e.g. Ponzetto & Strube (2007). Also, the analysis of other forms of structural information in Wikipedia apart from the outlinks might be promising.

Acknowledgements. The work described in this paper was partly funded by Deutsche Forschungsgemeinschaft (DFG), Project DIANA-Summ, STR 545/2-1,2, and by the Klaus Tschira Foundation. We thank our annotators Ganna Syrota and Alessandra Moschetti. We are also grateful to the anonymous LREC reviewers for valuable comments and suggestions.

References

- Clark, Herbert H. (1975). Bridging. In *Proceedings of the Conference on Theoretical Issues in Natural Language Processing*, Cambridge, Mass., June 1975, pp. 169–174.
- Fellbaum, Christiane (Ed.) (1998). *WordNet: An Electronic Lexical Database*. Cambridge, Mass.: MIT Press.
- Fleiss, Joseph L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382.
- Gupta, Surabhi, Matthew Purver & Dan Jurafsky (2007). Disambiguating between generic and referential “you” in dialog. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*, Prague, Czech Republic, June 23–30, 2007, pp. 105–108.
- Hahn, Udo, Michael Strube & Katja Markert (1996). Bridging textual ellipses. In *Proceedings of the 16th International Conference on Computational Linguistics*, Copenhagen, Denmark, 5–9 August 1996, Vol. 1, pp. 496–501.
- Janin, Adam, Don Baron, Jane Edwards, Dan Ellis, David Gelbart, Nelson Morgan, Barbara Peskin, Thilo Pfau, Elizabeth Shriberg, Andreas Stolcke & Chuck Wooters (2003). The ICSI Meeting Corpus. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Hong Kong, pp. 364–367.

Antecedent	Bridging Anaphor	Antecedent Entry Found?	Bridging Anaphor Found?	
			In Page Text	In Outlinks
microphone	batteries	yes	yes	no
data collection	data	no	-	-
neural net	training	yes	yes	no
experiment	result	yes	yes	no
table	column	no	-	-
easy problem	answer	yes	yes	no
question	answer	yes	yes	yes

Table 5: Results for Automatic Extraction from Wikipedia.

- Mihalcea, Rada & Andras Csomai (2007). Wikify!: linking documents to encyclopedic knowledge. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management (CIKM '07)*, pp. 233–242.
- Müller, Christoph (2007). Resolving *It*, *This*, and *That* in unrestricted multi-party dialog. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, Prague, Czech Republic, June 23–30, 2007, pp. 816–823.
- Müller, Christoph & Michael Strube (2006). Multi-level annotation of linguistic data with MMAX2. In Sabine Braun, Kurt Kohn & Joybrato Mukherjee (Eds.), *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*. Frankfurt a.M., Germany: Peter Lang.
- Nissim, Malvina, Shipra Dingare, Jean Carletta & Mark Steedman (2004). An annotation scheme for information status in dialogue. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, Lisbon, Portugal, 26–28 May, 2004.
- Poesio, Massimo, Renata Vieira & Simone Teufel (1997). Resolving bridging references in unrestricted text. In *Proceedings of the ACL Workshop on Operational Factors in Robust Anaphora Resolution*, pp. 1–6.
- Ponzetto, Simone Paolo & Michael Strube (2007). Knowledge derived from Wikipedia for computing semantic relatedness. *Journal of Artificial Intelligence Research*, 30:181–212.
- Prince, Ellen (1992). The ZPG letter: Subjects, Definiteness, and Information-status. In William. C. Mann & Sandra A. Thompson (Eds.), *Discourse Description: Diverse Linguistic analyses of a fund-raising text*, pp. 295–325. Philadelphia: John Benjamins.
- Schmid, Helmut (1997). Probabilistic part-of-speech tagging using decision trees. In Daniel Jones & Harold Somers (Eds.), *New Methods in Language Processing*, pp. 154–164. London, UK: UCL Press.
- Strube, Michael & Udo Hahn (1999). Functional centering: Grounding referential coherence in information structure. *Computational Linguistics*, 25(3):309–344.
- Zesch, Torsten, Christof Müller & Iryna Gurevych (2008). Extracting Lexical Semantic Knowledge from Wikipedia and Wiktionary. In *Proceedings of the Conference on Language Resources and Evaluation (LREC)*. To appear.