How bad is the problem of PP-attachment? A comparison of English, German and Swedish

Martin Volk
Stockholm University
Department of Linguistics
106 91 Stockholm, Sweden
volk@ling.su.se

Abstract

The correct attachment of prepositional phrases (PPs) is a central disambiguation problem in parsing natural languages. This paper compares the baseline situation in English, German and Swedish based on manual PP attachments in various treebanks for these languages. We argue that cross-language comparisons of the disambiguation results in previous research is impossible because of the different selection procedures when building the training and test sets. We perform uniform treebank queries and show that English has the highest noun attachment rate followed by Swedish and German. We also show that the high rate in English is dominated by the preposition *of*. From our study we derive a list of criteria for profiling data sets for PP attachment experiments.

1 Introduction

Any computer system for natural language processing has to struggle with the problem of ambiguities. If the system is meant to extract precise information from a text, these ambiguities must be resolved. One of the most frequent ambiguities arises from the attachment of prepositional phrases (PPs). Simply stated, a PP that follows a noun (in English, German or Swedish) can be attached to the noun or to the verb.

In the last decade various methods for the resolution of PP attachment ambiguities have been proposed. The seminal paper by (Hindle and Rooth, 1993) started a sequence of studies for English. We investigated similar methods for German (Volk, 2001; Volk, 2002). Recently other languages (such as Dutch (Vandeghinste, 2002) or Swedish (Aasa, 2004)) have followed.

In the PP attachment research for other languages there is often a comparison of the disambiguation accuracy with the English results. But are the results really comparable across languages? Are we starting from the same baseline when working on PP attachment in structurally similar languages like English, German and Swedish? Is the problem of PP attachment equally bad (equally frequent and of equal balance) for these three languages? These are the questions we will discuss in this paper.

In order to find answers to these questions we have taken a closer look at the training and test data used in various experiments. And we have queried the most important treebanks for the three languages under investigation.

2 Background

(Hindle and Rooth, 1993) did not have access to a large treebank. Therefore they proposed an unsupervised method for resolving PP attachment ambiguities. And they evaluated their method against 880 English triples verb-noun-preposition (V-N-P) which they had extracted from randomly selected, ambiguously located PPs in a corpus. For example, the sentence *"Timex had requested duty-free treatment for many types of watches"* results in the V-N-P triple *(request, treatment, for)*. These triples were manually annotated by both authors with either noun or verb attachment based on the complete sentence context. Interestingly, 586 of these triples (67%) were judged as noun attachments and only 33% as verb attachments. And (Hindle and Rooth, 1993) reported on 80% attachment accuracy, an improvement of 13% over the baseline (i.e. guessing noun attachment in all

Proceedings of the Third ACL-SIGSEM Workshop on Prepositions, pages 81–88,
Trento, Italy, April 2006.

cases).

A year later (Ratnaparkhi et al., 1994) published a supervised approach to the PP attachment problem. They had extracted quadruples V-N-P-N[1] (plus the accompanying attachment decision) from both an IBM computer manuals treebank (about 9000 tuples) and from the Wall Street Journal (WSJ) section of the Penn treebank (about 24,000 tuples). The latter tuple set has been reused by subsequent research, so let us focus on this one.[2] (Ratnaparkhi et al., 1994) used 20,801 tuples for training and 3097 tuples for evaluation. They reported on 81.6% correct attachments.

But have they solved the same problem as (Hindle and Rooth, 1993)? What was the initial bias towards noun attachment in their data? It turns out that their training set (the 20,801 tuples) contains only 52% noun attachments, while their test set (the 3097 tuples) contains 59% noun attachments. The difference in noun attachments between these two sets is striking, but (Ratnaparkhi et al., 1994) do not discuss this (and we also do not have an explanation for this). But it makes obvious that (Ratnaparkhi et al., 1994) were tackling a problem different from (Hindle and Rooth, 1993) given the fact that their baseline was at 59% guessing noun attachment (rather than 67% in the Hindle and Rooth experiments).[3]

Of course, the baseline is not a direct indicator of the difficulty of the disambiguation task. We may construct (artificial) cases with low baselines and a simple distribution of PP attachment tendencies. For example, we may construct the case that a language has 100 different prepositions, where 50 prepositions always introduce noun attachments, and the other 50 prepositions always require verb attachments. If we also assume that both groups occur with the same frequency, we have a 50% baseline but still a trivial disambiguation task.

But in reality the baseline puts the disambiguation result into perspective. If, for instance, the baseline is 60% and the disambiguation result is 80% correct attachments, then we will claim that our disambiguation procedure is useful. Whereas if we have a baseline of 80% and the disambiguation result is 75%, then the procedure can be discarded.

So what are the baselines reported for other languages? And is it possible to use the same extraction mechanisms for V-N-P-N tuples in order to come to comparable baselines?

We did an in-depth study on German PP attachment (Volk, 2001). We compiled our own treebank by annotating 3000 sentences from the weekly computer journal ComputerZeitung. We had first annotated a larger number of subsequent sentences with Part-of-Speech tags, and based on these PoS tags, we selected 3000 sentences that contained at least one full verb plus the sequence of a noun followed by a preposition. After annotating the 3000 sentences with complete syntax trees we used a Prolog program to extract V-N-P-N tuples with the accompanying attachment decisions. This lead to 4562 tuples out of which 61% were marked as noun attachments. We used the same procedure to extract tuples from the first 10,000 sentences of the NEGRA treebank. This resulted in 6064 tuples with 56% noun attachment (for a detailed overview see (Volk, 2001) p. 86). Again we observe a substantial difference in the baseline.

When our student Jörgen Aasa worked on replicating our German experiments for Swedish, he used a Swedish treebank from the 1980s for the extraction of test data. He extracted V-N-P-N tuples from SynTag, a treebank with 5100 newspaper sentences built by (Järborg, 1986). And Aasa was able to extract 2893 tuples out of which 73.8% were marked as noun attachments (Aasa, 2004) (p. 25). This was a surprisingly high figure, and we wondered whether this indicated a tendency in Swedish to avoid the PP in the ambiguous position unless it was to be attached to the noun. But again the extraction process was done with a special purpose extraction program whose correctness was hard to verify.

3 Querying Treebanks with TIGER-Search

We therefore decided to check the attachment tendencies of PPs in various treebanks for the three languages in question with the same tool and with queries that are as uniform as possible.

For English we used the WSJ section of the Penn Treebank, for German we used our own ComputerZeitung treebank (3000 sentences), the

[1]The V-N-P-N quadruples also contain the head noun of the NP within the PP.

[2]The Ratnaparkhi training and test sets were later distributed together with a development set of 4039 V-N-P-N tuples.

[3]It should be noted that important subsequent research, e.g. by (Collins and Brooks, 1995; Stetina and Nagao, 1997), used the Ratnaparkhi data sets and thus allowed for good comparability.

NEGRA treebank (10,000 sentences) and the recently released version of the TIGER treebank (50,000 sentences). For Swedish we used the SynTag treebank mentioned above and one section of the Talbanken treebank (6100 sentences). All these treebanks consist of constituent structure trees, and they are in representation formats which allow them to be loaded into TIGER-Search. This enables us to query them all in similar manners and to get a fairer comparison of the attachment tendencies.

TIGER-Search is a powerful treebank query tool developed at the University of Stuttgart (König and Lezius, 2002). Its query language allows for feature-value descriptions of syntax graphs. It is similar in expressiveness to tgrep (Rohde, 2005) but it comes with graphical output and highlighting of the syntax trees plus some nice statistics functions.

Our experiments for determining attachment tendencies proceed along the following lines. For each treebank we first query for all sequences of a noun immediately followed by a PP (henceforth noun+PP sequences). The dot being the precedence operator, we use the query:

```
[pos="NN"] . [cat="PP"]
```

This query will match twice in the tree in figure 1. It gives us the frequency of all ambiguously located PP. We disregard the fact that in certain clause positions a PP in such a sequence cannot be verb-attached and is thus not ambiguous. For example, an English noun+PP sequence in subject position is not ambiguous with respect to PP attachment since the PP cannot attach to the verb. Similar restrictions apply to German and Swedish.

In order to determine how many of these sequences are annotated as noun attachments, we query for noun phrases that contain both a noun and an immediately following PP. This query will look like:

```
#np_mum:[cat="NP"]  >
     #np_child:[cat="NP"] &
#np_mum   >  #pp:[cat="PP"] &
#np_child >* #noun:[pos="NN"] &
#noun . #pp
```

All strings starting with # are variables and the $>$ symbol is the dominance operator. So, this query says: Search for an NP (and call it np_mum) that immediately dominates another NP (np_child) AND that immediately dominates a PP, AND the np_child dominates a noun which is immediately followed by the PP.

This query presupposes that a PP which is attached to a noun is actually annotated with the structure (NP (NP (... N)) (PP)) which is true for the Penn treebank (compare to the tree in figure 1). But the German treebanks represent this type of attachment rather as (NP (... N) (PP)) which means that the query needs to be adapted accordingly.[4]

Such queries give us the frequency of all noun+PP sequences and the frequency of all such sequences with noun attachments. These frequencies allow us to calculate the noun attachment rate (NAR) in our treebanks.

$$NAR = \frac{freq(noun + PP, noun\_attachm)}{freq(noun + PP)}$$

We assume that all PPs in noun+PP sequences which are not attached to a noun are attached to a verb. This means we ignore the very few cases of such PPs that might be attached to adjectives (as for instance the second PP in *"due for revision in 1990"*).

Different annotation schemes require modifications to these basic queries, and different noun classes (regular nouns, proper names, deverbal nouns etc.) allow for a more detailed investigation. We now present the results for each language in turn.

3.1 Results for English

We used sections 0 to 12 of the WSJ part of the Penn Treebank (Marcus et al., 1993) with a total of 24,618 sentences for our experiments. Our start query reveals that an ambiguously located PP (i.e. a noun+PP sequence) occurs in 13,191 (54%) of these sentences, and it occurs a total of 20,858 times (a rate of 0.84 occurrences per sentences with respect to all sentences in the treebank).

Searching for noun attachments with the second query described in section 3 we learn that 15,273 noun+PP sequences are annotated as noun attachments. And we catch another 547 noun attachments if we query for noun phrases that contain two PPs in sequence.[5] In these cases the second PP is also attached to a noun, although not

[4] There are a few occurrences of this latter structure in the Penn Treebank which should probably count as annotation errors.

[5] See (Merlo et al., 1997) for a discussion of these cases and an approach in automatically disambiguating them.

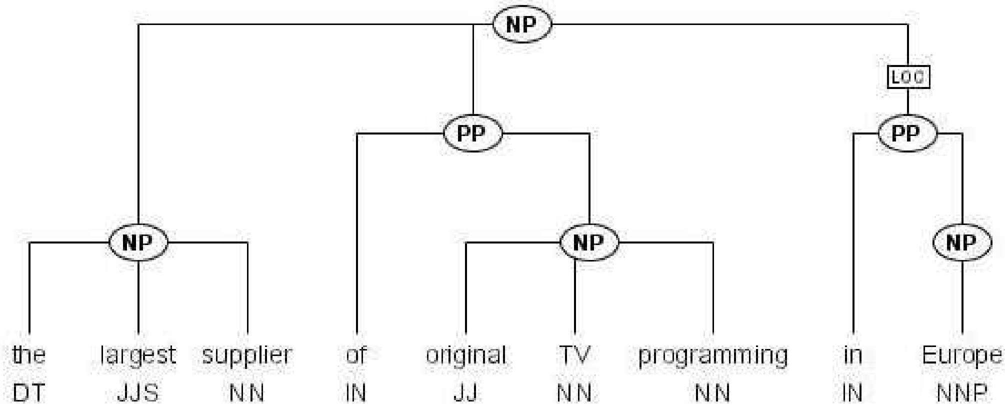

Figure 1: Noun phrase tree from the Penn Treebank

to the noun immediately preceding it (as for example in the tree in figure 1). With some similar queries we located another 110 cases of noun attachments (most of which are probably annotation errors if the annotation guidelines are applied strictly). This means that we found a total of 15,930 cases of noun attachment which corresponds to a noun attachment rate of 76.4% (by comparison to the 20,858 occurrences).

This is a surprisingly high number. Neither (Hindle and Rooth, 1993) with 67% nor (Ratnaparkhi et al., 1994) with 59% noun attachment were anywhere close to this figure. What have we done differently?

One aspect is that we only queried for singular nouns (NN) in the Penn Treebank where plural nouns (NNS) and proper names (NNP and NNPS) have separate PoS tags. Using analogous queries for plural nouns we found that they exhibit a NAR of 71.7%. Whereas the queries for proper names (singular and plural names taken together) account for a NAR of 54.5%.

Another reason for the discrepancy in the NAR between Ratnaparkhi's data and our calculations certainly comes from the fact that we queried for all sequences noun+PP as possibly ambiguous whereas they looked only at such sequences within verb phrases. But since we will do the same in both German and Swedish, this is still worthwhile.

3.2 Results for German

The three German treebanks which we investigate are all annotated in more or less the same manner, i.e. according to the NEGRA guidelines which were slightly refined for the TIGER project. This enabled us to use the same set of queries for all three of them. Since the German guidelines distinguish between node labels for coordinated phrases (e.g. CNP and CPP) and non-coordinated phrases (e.g. NP and PP), these distinctions needed to be taken into account. Table 1 summarizes the results.

	CZ	NEGRA	TIGER
size	3000	10,000	50,000
noun+PP seq	4355	6,938	39,634
occur rate	1.4	0.7	0.8
noun attachm	2743	4102	23,969
NAR	63.0%	59.1%	60.5%

Table 1: Results for the German treebanks

Our own ComputerZeitung treebank (CZ) has a much higher occurrence rate of ambiguously located PPs because the sentences were preselected for this phenomenon. The general NEGRA and TIGER treebanks have an occurrence rate that is similar to English (0.8). The NAR varies between 59.1% for the NEGRA treebank and 63.0% for the CZ treebank for regular nouns.

The German annotation also distinguishes between regular nouns and proper names. The proper names show a much lower noun attachment rate than the regular nouns. The NAR in the CZ treebank is 22%, in the NEGRA treebank it is 20%, and in the TIGER treebank it is only 17%. Here we suspect that the difference between the CZ and the other treebanks is based on the different text types. The computer journal CZ contains more person names with affiliation (e.g. *Stan Sugarman von der Firma Telemedia*) and more company names with location (e.g. *Aviso aus Finn-*

land) than a regular newspaper (that was used in the NEGRA and TIGER corpora).

As mentioned above, our previous experiments in (Volk, 2001) were based on sets of extracted tuples from both the CZ and NEGRA treebanks. Our extracted data set from the CZ treebank had a noun attachment rate of 61%, and the one from the NEGRA treebank had a noun attachment rate of 56%.

So why are our new results based on TIGER-Search queries two to three percents higher? The main reason is that our old data sets included proper names (with their low noun attachment rate). But our extraction procedure comprised also a number of other idiosyncracies. In an attempt to harvest as many interesting V-N-P-N tuples as possible from our treebanks we exploited coordinated phrases and pronominal PPs. Some examples:

1. If the PP was preceded by a coordinated noun phrase, we created as many tuples as there were head nouns in the coordination. For example, the phrase *"den Austausch und die gemeinsame Nutzung von Daten ... ermöglichen"* leads to the tuples (*ermöglichen, Austausch, von, Daten*) and (*ermöglichen, Nutzung, von, Daten*) both with the decision 'noun attachment'.

2. If the PP was introduced by coordinated prepositions (e.g. *Die Argumente für oder gegen den Netzwerkcomputer*), we created as many tuples as there were prepositions.

3. If the verb group consists of coordinated verbs (e.g. *Infos für Online-Dienste aufbereiten und gestalten*), we created as many tuples as there were verbs.

4. We regarded pronominal adverbs (*darin, dazu, hierüber*, etc.) and reciprocal pronouns (*miteinander, untereinander, voneinander*, etc.) as equivalent to PPs and created tuples when such pronominals appeared immediately after a noun. See (Volk, 2003) for a more detailed discussion of these pronouns.

3.3 Results for Swedish

Currently there is no large-scale Swedish treebank available. But there are some smaller treebanks from the 80s which have recently been converted to TIGER-XML so that they can also be queried with TIGER-Search.

SynTag (Järborg, 1986) is a treebank consisting of around 5100 sentences. Its conversion to TIGER-XML is documented in (Hagström, 2004). The treebank focuses on predicate-argument structures and some grammatical functions such as subject, head and adverbials. It is thus different from the constituent structures that we find in the Penn treebank or the German treebanks. We had to adapt our queries accordingly. Since prepositional phrases are not marked as such, we need to query for constituents (marked as subject, as adverbial or simply as argument) that start with a preposition. This results in a noun attachment rate of 73% (which is very close to the rate reported by (Aasa, 2004)). Again this does not include proper names which have a NAR of 44% in SynTag.

Let us compare these results to the second Swedish treebank, Talbanken (first described by (Telemann, 1974)). Talbanken was a remarkable achievement in the 80s as it comes with two written language parts (with a total of more than 10,000 sentences from student essays and from newspapers) and two spoken language parts (with another 10,000 trees from interviews and conversations). We concentrated on the 6100 trees from the written part taken from newspaper texts.

The occurrence rate in Talbanken is 0.76 (4658 noun+PP sequences in 6100 sentences), which is similar to the rates observed for English and German. The occurrence rate in SynTag is higher 0.93 (4737 noun+PP sequences in 5114 sentences).

Talbanken (in its converted form) is annotated with constituent structure labels (NP, PP, VP etc.) and also distinguishes coordinated phrases (CNP, CPP, CVP etc.). The queries for determining the noun attachment rate can thus be similar to the queries over the German treebanks. In addition, Talbanken comes with a rich set of grammatical features as edge labels (e.g. there are different labels for logical subject, dummy subject and other subject).

We found that the NAR for regular nouns in Talbanken is 60.5%. Talbanken distinguishes between regular nouns, deverbal nouns (often with the derivation suffix *-ing: tjänstgöring, utbildning, övning*) and deadjectival nouns (mostly with the derivation suffix *-het: skyldighet, snabbhet, verksamhet*). Not surprisingly, these special nouns have higher NARs than the regular nouns. The deadjectival nouns have a NAR of 69.5%, and the deverbal nouns even have a NAR of 77%. Taken together (i.e. regarding all regular, deadjectival and deverbal nouns) this results in a NAR of 64%.

Thus, the NARs which we obtain from the two Swedish treebanks (SynTag 73% and Talbanken 64%) differ drastically. It is unclear what this difference depends on. The text genre (newspapers) is the same in both cases. We have noticed that SynTag contains a number of annotation errors, but we don't see that these errors favor noun attachment of PPs in a systematic way. One aspect might be the annotation decision in Talbanken to annotate PPs in light verb constructions.

These are disturbing cases where the PP is a child node of the sentence node S (which means that it is interpreted as a verb attachment) with the edge label OA (*objektadverbial*). Nivre (2005, personal communication) pointed out that "OA is what some theoreticians would call a 'prepositional object' or a 'PP complement', i.e. a complement of the verb that semantically is close to an object but which is realized as a prepositional phrase." In our judgement many of those cases should be noun attachments (and thus be a child of an NP).

For example, we looked at *förutsättning för* (= prerequisite for) which occurs 14 times, out of which 2 are annotated as OO (Other object) + OA, 11 are annotated as noun attachments, and 1 is erroneously annotated. If we compare that to *betydelse för* (= significance for) which occurs 16 times out of which 13 are annotated as OO+OA and 3 are annotated as noun attachments, we wonder.

First, it is obvious that there are inconsistencies in the treebank. We cannot see any reason why the 2 cases of *förutsättning för* are annotated differently than the other 11 cases. The verbs do not justify these discrepancies. For example, we have *skapa* (= to create) with the verb attachments and *försvinna* (= to disappear) with the noun attachment cases. And we find *ge* (= to give) on both sides.

Second, we find it hard to follow the argument that the tendency for *betydelse för* is stronger for the OO+OA than for *förutsättning för*. It might be based on the fact that *betydelse för* is often used with the verb *ha* (= to have) and thus may count as a light verb construction with a verb group consisting of both *ha* plus *betydelse* and the *för*-PP being interpreted as an object of this complex verb group.

Third, unfortunately not all cases of PPs annotated as *objektadverbial* can be regarded as noun attachments. But after having looked at some 70 occurrences of such PPs immediately following a noun, we estimate that around 30% should be noun attachments.

Concluding our observations on Swedish let us mention that the very few cases of proper names in Talbanken have a NAR of 24%.

4 Comparison of the results

For English we have computed a NAR of 76.4% based on the Penn Treebank, for German we found NARs between 59% and 63% based on three treebanks, and for Swedish we determined a puzzling difference between 73% NAR in SynTag and 64% NAR in Talbanken. So, why is the tendency of a PP to attach to a preceding noun stronger in English than in Swedish which in turn shows a stronger tendency than German?

For English the answer is very clear. The strong NAR is solely based on the dominance of the preposition *of*. In our section of the Penn Treebank we found 20,858 noun+PP sequences. Out of these, 8412 (40% !!) were PPs with the preposition *of*. And 99% of all *of*-PPs are noun attachments. So, the preposition *of* dominates the English NAR to the point that it should be treated separately.[6]

The Ratnaparkhi data sets (described above in section 2) contain 30% tuples with the preposition *of* in the test set and 27% *of*-tuples in the training set. The higher percentage of *of*-tuples in the test set may partially explain the higher NAR of 59% (vs. 52% in the training set).

The dominance of *of*-tuples may also explain the relatively high NAR for proper names in English (54.5%) in comparison to 17% - 22% in German and similar figures for the Swedish Talbanken corpus. The Penn Treebank represents names that contain a PP (e.g. *District of Columbia, American Association of Individual Investors*) with a regular phrase structure. It turns out that 861 (35%) of the 2449 sequences 'proper name followed by PP' are based on *of*-PPs. The dominance becomes even more obvious if we consider that the following

[6]This is actually what has been done in some research on English PP attachment disambiguation. (Ratnaparkhi, 1998) first assumes noun attachment for all *of*-PPs and then applies his disambiguation methods to all remaining PPs.

prepositions on the frequency ranks are *in* (with only 485 occurrences) and *for* (246 occurrences).

The dominance of the preposition *of* is so strong in English that we will get a totally different picture of attachment preferences if we omit *of*-PPs. The Ratnaparkhi training set without *of*-tuples is left with a NAR of 35% (!) and the test set has a NAR of 42%. In other words, English has a clear tendency of attaching PPs to verbs if we ignore the dominating *of*-PPs.

Neither German nor Swedish has such a dominating preposition. There are, of course, prepositions in both languages that exhibit a clear tendency towards noun attachment or verb attachment. But they are not as frequent as the preposition *of* in English. For example, clear temporal prepositions like German *seit* (= since) are much more likely as verb attachments.

Closest to the English *of* is the Swedish preposition *av* which has a NAR of 88% in the Talbanken corpus. But its overall frequency does not dominate the Swedish ranking. The most frequent preposition in ambiguous positions is *i* (frequency: 651 and NAR: 53%) followed by *av* (frequency: 564; NAR: 88%) and *för* (frequency: 460; NAR: 42%).

5 Conclusion

The most important conclusion to be drawn from the above experiments and observations is the importance of profiling the data sets when working and reporting on PP attachment experiments. The profile should certainly answer the following questions:

1. What types of nouns where used when the tuples were extracted? (regular nouns, proper names, deverbal nouns, etc.)

2. Are there prepositions which dominate in frequency and attachment rate (like the English preposition *of*)? If so, how does the data set look like without these dominating prepositions?

3. What types of prepositions where regarded? (regular prepositions, contracted prepositions (e.g. in German *am, im, zur*), derived prepositions (e.g. English prepositions derived from gerund verb forms *following, including, pending*) etc.)

4. Is the extraction procedure restricted to noun+PP sequences in the verb phrase, or does it consider all such sequences?

5. What is the noun attachment rate in the data set?

In order to find dominating prepositions we suggest a data profiling that includes the frequency and NARs of all prepositions in the data set. This will also give an overall picture of the number of prepositions involved.

Our experiments have also shown the advantages of large treebanks for comparative linguistic studies. Such treebanks are even more valuable if they come in the same representation schema (e.g. TIGER-XML) so that they can be queried with the same tools. TIGER-Search has proven to be a suitable treebank query tool for our experiments although its statistics function broke down on some frequency counts we tried on large treebanks. For example, it was not possible to get a list of all prepositions with occurrence frequencies from a 50,000 sentence treebank.

Another item on our TIGER-Search wish list is a batch mode so that we could run a set of queries and obtain a list of frequencies. Currently we have to trigger each query manually and copy the frequency results manually to an Excel file.

Other than that, TIGER-Search is a wonderful tool which allows for quick sanity checks of the queries with the help of the highlighted tree structure displays in its GUI.

We have compared noun attachment rates in English, German and Swedish over treebanks from various sources and with various annotation schemes. Of course, the results would be even better comparable if the treebanks were built on the same translated texts, i.e. on parallel corpora. Currently, there are no large parallel treebanks available. But our group works on such a parallel treebank for English, German and Swedish. Design decisions and first results were reported in (Volk and Samuelsson, 2004) and (Samuelsson and Volk, 2005). We believe that such parallel treebanks will allow a more focused and more detailed comparison of phenomena across languages.

6 Acknowledgements

We would like to thank Jörgen Aasa for discussions on PP attachment in Swedish, and Joakim

Nivre, Johan Hall, Jens Nilsson at Växjö University for making the Swedish Talbanken treebank available. We also thank the anonymous reviewers for their discerning comments.

References

Jörgen Aasa. 2004. Unsupervised resolution of PP attachment ambiguities in Swedish. Master's thesis, Stockholm University. Combined C/D level thesis.

Michael Collins and James Brooks. 1995. Prepositional phrase attachment through a backed-off model. In *Proc. of the Third Workshop on Very Large Corpora*.

Bo Hagström. 2004. A TIGER-XML version of SynTag. Master's thesis, Stockhom University.

D. Hindle and M. Rooth. 1993. Structural ambiguity and lexical relations. *Computational Linguistics*, 19(1):103–120.

Jerker Järborg. 1986. SynTag Dokumentation. Manual för SynTaggning. Technical report, Department of Swedish, Göteborg University.

Esther König and Wolfgang Lezius. 2002. The TIGER language - a description language for syntax graphs. Part 1: User's guidelines. Technical report.

Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn treebank. *Computational Linguistics*, 19(2):313–330.

P. Merlo, M.W. Crocker, and C. Berthouzoz. 1997. Attaching multiple prepositional phrases: generalized backed-off estimation. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing. Brown University, RI*.

A. Ratnaparkhi, J. Reynar, and S. Roukos. 1994. A maximum entropy model for prepositional phrase attachment. In *Proceedings of the ARPA Workshop on Human Language Technology*, Plainsboro, NJ, March.

Adwait Ratnaparkhi. 1998. Statistical models for unsupervised prepositional phrase attachment. In *Proceedings of COLING-ACL-98*, Montreal.

Douglas L. T. Rohde, 2005. *TGrep2 User Manual*. MIT. Available from http://tedlab.mit.edu/∼dr/Tgrep2/.

Yvonne Samuelsson and Martin Volk. 2005. Presentation and representation of parallel treebanks. In *Proc. of the Treebank-Workshop at Nodalida*, Joensuu, May.

J. Stetina and M. Nagao. 1997. Corpus-based PP attachment ambiguity resolution with a semantic dictionary. In J. Zhou and K. Church, editors, *Proc. of the 5th Workshop on Very Large Corpora*, pages 66–80, Beijing and Hongkong.

Ulf Telemann. 1974. *Manual För Grammatisk Beskrivning Av Talad Och Skriven Svenska*. Inst. för nordiska språk, Lund.

Vincent Vandeghinste. 2002. Resolving PP attachment ambiguities using the WWW (abstract). In *Computational Linguistics in the Netherlands*, Groningen.

Martin Volk and Yvonne Samuelsson. 2004. Bootstrapping parallel treebanks. In *Proc. of Workshop on Linguistically Interpreted Corpora (LINC) at COLING*, Geneva.

Martin Volk. 2001. *The automatic resolution of prepositional phrase attachment ambiguities in German*. Habilitationsschrift, University of Zurich.

Martin Volk. 2002. Combining unsupervised and supervised methods for PP attachment disambiguation. In *Proc. of COLING-2002*, Taipeh.

Martin Volk. 2003. German prepositions and their kin. a survey with respect to the resolution of PP attachment ambiguities. In *Proc. of ACL-SIGSEM Workshop: The Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*, pages 77–88, Toulouse, France, September. IRIT.