Developments in the TIGER Annotation Scheme and their Realization in the Corpus

Sabine Brants*, **Silvia Hansen***

* Computational Linguistics, Saarland University
Postfach 151150, 66041 Saarbrücken, Germany
{sabine, hansen}@coli.uni-sb.de

Abstract

This paper presents the annotation of the German TIGER Treebank. First, issues concerning the annotation, representation as well as querying of the treebank are discussed. Within this context, the annotation tool ANNOTATE, the export and XML formats of the TIGER Treebank and the TIGER search tool are briefly introduced. Secondly, the developments of the TIGER annotation scheme and their realization in the corpus are introduced focussing on the differences between the underlying NEGRA annotation scheme and the further developed TIGER annotation scheme. The main differences are concerned with verb-subcategorization, coordination, appositions and parentheses as well as proper nouns. Thirdly, the annotation scheme is assessed through an evaluation and a problem discussion of the above mentioned changes. For this purpose, inter-annotator agreement in the TIGER project has been analyzed focussing on exactly these changes. This analysis shows where the annotators' decision problems are. These difficulties are discussed in greater detail on the basis of annotation examples. The paper concludes with some suggestions for the improvement of the TIGER annotation scheme.

1. Introduction

There has been an increasing interest in recent years in the enrichment of natural language corpora in terms of annotation with syntactic information. One of the best known treebanks is the English Penn Treebank (Marcus et al., 1994) - but comparable treebanks exist for English, such as the Susanne Corpus (Sampson, 1995), the Lancaster Parsed Corpus (Leech, 1992) and the British part of the International Corpus of English (Greenbaum, 1996), as well as for other languages, such as the Prague Dependency Treebank for Czech (Hajic, 1999). Recently, treebank projects for other languages have come to life as well, e.g. for French (Abeillé et al., 2000b), Italian (Bosco et al., 2000), Spanish (Moreno et al., 2000), Turkish (Oflazer et al., 1999) and Russian (Boguslavsky et al., 2000). More initiatives for linguistically interpreted corpora can be found in (Uszkoreit et al., 1999) and (Abeillé et al., 2000a).

For German there are three syntactically annotated corpora: the Verbmobil Corpus (Wahlster, 2000), the NEGRA Corpus (Skut et al., 1998) and the TIGER Treebank (Dipper et al., 2001). Since the Verbmobil Corpus is rather restricted in its domains (i.e. spontaneous speech for the appointment negotiation domain) and the NEGRA Corpus in its size (20,000 syntactically annotated sentences), the annotation of the TIGER Treebank as a comprehensive resource for German linguists was more than overdue.

2. Annotation of a German Treebank

The basis of the TIGER Treebank are texts from the German newspaper 'Frankfurter Rundschau'. The linguistic annotation of each sentence in the TIGER Treebank is represented in terminal nodes (for parts-of-speech), nonterminal nodes (for phrase categories) and edges (for syntactic functions). Furthermore, a supplementary annotation on the word level is used to encode information on lemmata and morphology. Secondary edges are used to encode coordination information.

In order to provide accurate results, each sentence of the TIGER Treebank is annotated independently by two annotators followed by a consistency check. After the first project phase, the TIGER Treebank consists of approximately 50,000 syntactically annotated sentences. The aim of the second project phase is to extend this amount to about 80,000 sentences.

2.1. Corpus annotation

The annotation is done with a system called ANNOTATE (under development in the NEGRA and TIGER projects), which allows interactive parsing, i.e. a parser carries out a shallow parse and a human may disambiguate or correct the proposed parse (Plaehn and Brants, 2000). ANNOTATE uses the TnT tagger for part-of-speech tagging (Brants, 2000) and Cascaded Markov Models for the analysis of phrase categories (node labels) and syntactic functions (edge labels) (Brants, 1999). For the analysis of lemmata and morphological tags, ANNOTATE is interleaved with a tool called TigerMorph. For part-of-speech tagging as well as morphological annotation the Stuttgart-Tübingen-Tagset (Schiller et al., 1999) is used in a slightly modified version ((Kramp and Preis, 2000) and (Smith and Eisenberg, 2000)).

In addition to the interactive annotation combining automatic probabilistic parsing and human intervention, the TIGER corpus is parsed on the basis of LFG (using the Xerox Linguistic Environment) followed by semi-automatic disambiguation and the conversion of the LFG annotation into the TIGER format (Zinsmeister et al., 2001).

2.2. Corpus representation

The annotated sentences are stored and maintained in a MySQL database; in addition to that other output formats are possible as well. The transformation of the database entries into the TIGER export format is a straightforward step, since words, morphological tags, terminal nodes, nonterminal nodes and edges can be exported to a table stored

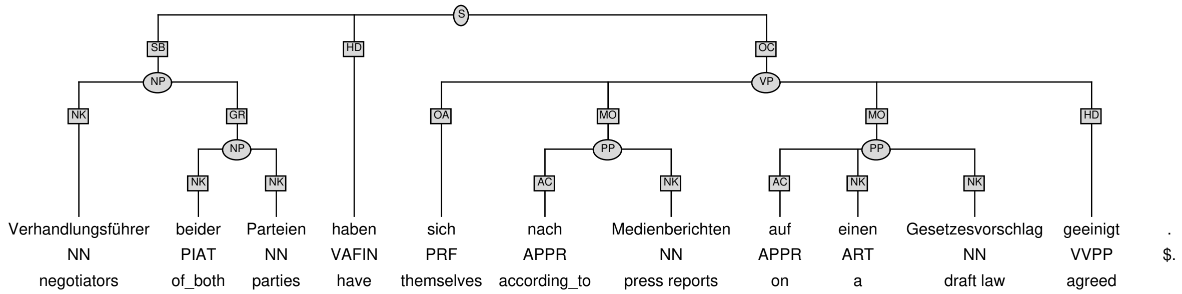

Figure 1: Annotation of PPs in NEGRA: MO

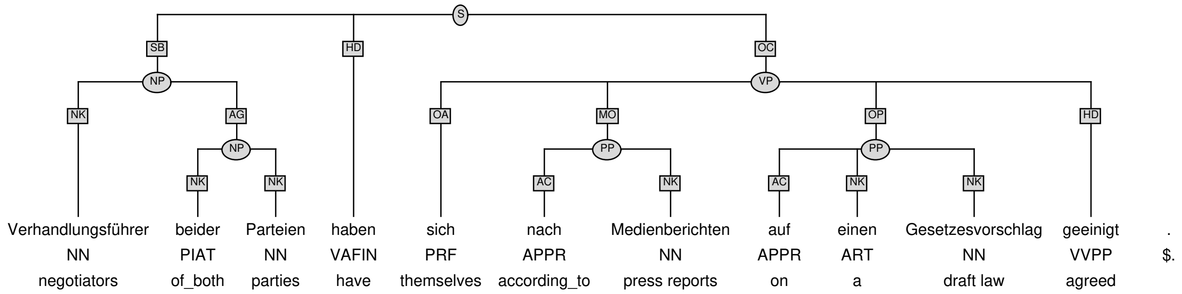

Figure 2: Annotation of PPs in TIGER: MO and OP

in a line-oriented and ASCII-based format (Brants, 1997). Sentence boundaries can be identified through sentence start and end tags. Furthermore, information on sentence origins, editors and used tags is stored at the beginning of each export file. The major advantage of the export format, which has been developed in the NEGRA project, is that it is easily readable for humans as well as easily processable for machines.

On the basis of the export format, the TIGER treebank can be transferred into the TIGER XML format (Lezius et al., 2002a). A TIGER XML document is split up into header and body. In the corpus header, meta-information on the corpus (such as corpus name, date, author, etc.) and an annotation grammar can be found. The annotation grammar consists of a declaration of the tags used for morphology, part-of-speech, non-terminal nodes and edges. For the corpus body, directed acyclic graphs are used as the underlying data model to encode the linguistic annotation. Thus, words, part-of-speech tags as well as morphological tags occur as attributes of the element 'terminal', whereas non-terminals are encoded through an additional element called 'nonterminal' refering to the corresponding terminal ID. Secondary edges are encoded explicitly as well. Through the use of XML, the TIGER Treebank is exchangable and usable with a large range of tools.

2.3. Corpus querying

The TIGER Treebank can be searched using the TIGER search tool (Lezius and König, 2000). This system is specialized in querying syntactically annotated corpora, since it provides an indexing mechanism for the rather complex representation of non-terminal nodes and edges. The preparatory steps, such as format conversions into TIGER XML, are carried out by the TIGER registry tool (Lezius et al., 2002b), which provides a filter for the import of the most popular treebank formats. Thus, other syntactically annotated corpora can be processed as well. The TIGER search tool operates on the basis of Boolean expressions, relations (e.g. precedence, dominance, etc.) and the combination of both (allowing restricted Boolean expressions over relations) (Lezius et al., 2002c). The matches of a query are displayed in form of trees, but can also be exported to TIGER XML format, which allows further processing through XSLT stylesheets.

3. Improvements in the TIGER annotation scheme

The basis for the annotation in TIGER is the annotation scheme that was developed in the NEGRA project (Brants et al., 1999). An important aspect of the work in the TIGER project is the extension of the annotation scheme. Although the NEGRA scheme covered a wide range of phenomena, there was still room for improvement in the linguistic adequacy of the annotation. In the following, the major changes that were made are discussed in detail. We also indicate problems and inconsistencies that arise from some of these changes.

3.1. Verb-subcategorization

In the NEGRA annotation scheme, it was not possible to distinguish prepositional phrases according to their functions; all PPs in sentences or verb phrases were unexceptionally marked with the label MO (modifier). In the TIGER project, two additional edge labels for PPs were introduced, namely prepositional objects (OP) and collocational verb constructions (CVC). The label OP is used

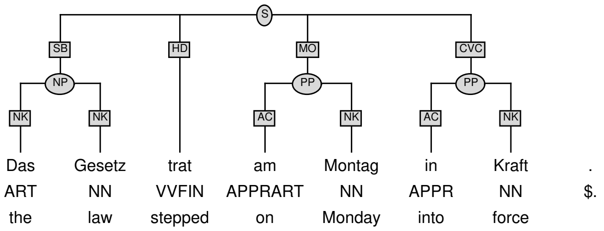

Figure 3: Annotation of PPs in TIGER: MO and CVC

for constructions like 'auf jemanden warten' ('to wait for somebody') where the preposition 'auf' ('on') has lost its lexical meaning. The different labels are illustrated in figures 1 to 3.

Figure 1 illustrates the fact that in the NEGRA annotation scheme, it was not possible to distinguish complements and adjuncts just by looking at the edge labels. In the TIGER annotation (figure 2), the functional difference between the first PP and the second PP is mirrored in the use of different edge labels: PP1 is functionally independent of the verb and serves as an adverbial; it is labeled with the old label MO. PP2 is a typical example for a prepositional object (OP) in German: the preposition 'auf' ('on') has lost its lexical meaning and has purely functional character.

The label CVC is reserved for verb + PP constructions where the main semantic information is contained in the noun of the PP, not in the verb (figure 3). This label can only be applied to a very limited class of verbs (semantically weak verbs with an originally directional or local meaning, e.g. 'stellen', 'kommen', etc. ('to put', 'to come')) in connection with an equally limited class of prepositions (mostly 'zu' ('to') and 'in' ('in')). An example for this is the German collocation 'in Kraft treten' (literally: ' to step into force', meaning: 'to take effect').

3.2. Coordination

Another important extension was made regarding the treatment of coordinated verb phrases and sentences. In NEGRA, arguments that have a syntactic role in both parts of a coordination, but that are only mentioned once, were structurally linked only to the nearest part of the coordination. In TIGER, these shared arguments are provided with secondary edges that represent the syntactic relationships of these arguments to the more distant verb conjuncts.

Figure 4 shows a sentence where the second verb in the coordination is intransitive (i.e. the second verb doesn't share the argument of the first verb), while figure 5 illustrates a sentence where the second verb is transitive and shares the direct object with the first verb. According to the NEGRA scheme, there is no structural difference in the annotation trees. Figure 6 shows the same sentence as figure 5, but annotated with the secondary edges that were introduced in TIGER. The sentence in figure 4 would be annotated in the same way in NEGRA and in TIGER. Thus, the TIGER scheme allows the differentiation between transitive and intransitive verbs in coordinations.

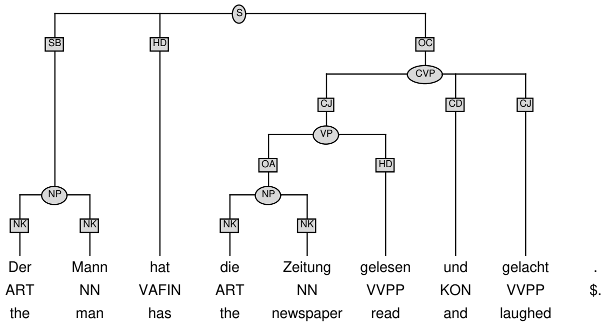

Figure 4: Coordination without shared arguments

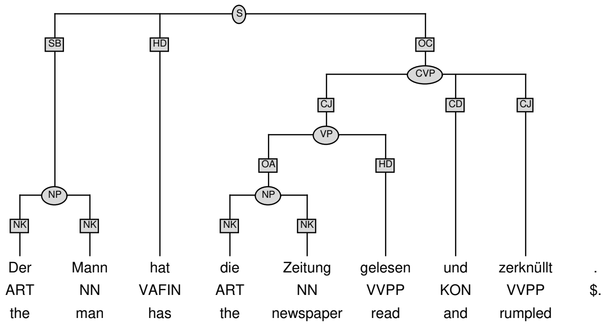

Figure 5: Coordination with shared argument in NEGRA

3.3. Appositions and parentheses

In NEGRA, the label APP (apposition) was used for many purposes. Not only was it used for 'true' appositions as in the following example:

- Dieter Schlenstedt, [der Präsident des ostdeutschen PEN]APP
- Dieter Schlenstedt, [the president of the East German PEN]APP

It also marked inserted phrases with parenthetical character which are mostly put in between brackets, for instance:

- Innenminister Otto Schily [(SPD)]APP

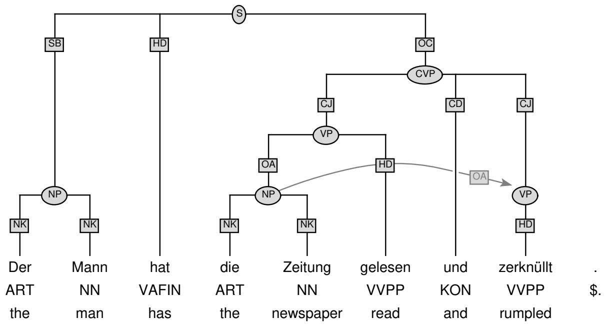

Figure 6: Coordination with shared argument in TIGER

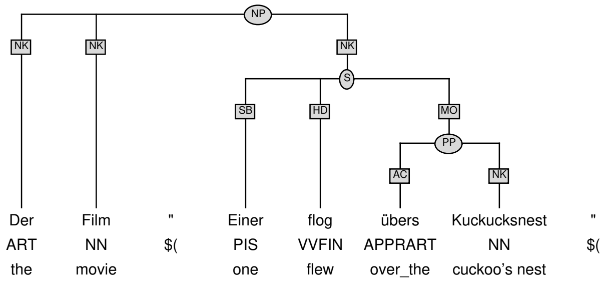

Figure 7: Treatment of structured proper nouns in NEGRA

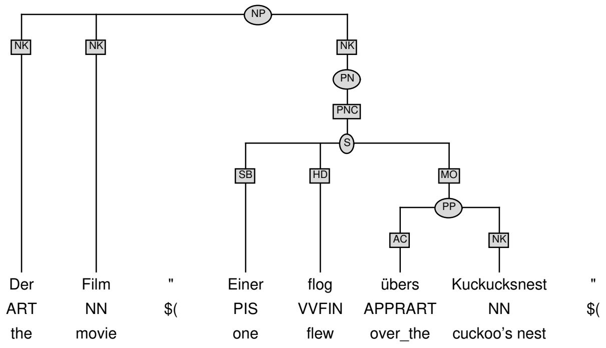

Figure 8: Annotation of structured proper nouns in TIGER

- Secretary of the Interior Otto Schily [(Social-Democrats)]APP

Cases like the last one are characterized by the lack of coreference between the inserted phrase and the main phrase, the presence of which is a prerequisite for true appositions; the label PAR (parenthesis) has been added to the annotation scheme to cover these cases more adequately:

- Innenminister Otto Schily [(SPD)]PAR

The annotation of 'true' appositions (i.e. with coreference) remains the same in TIGER.

3.4. Proper nouns

The label PN marks proper nouns, for instance 'George W. Bush'. The components of a proper noun are marked with the edge label PNC (proper noun component). In NEGRA, the label PN was also used for multitoken company names, newspaper names (e.g. 'The New York Times') etc. In the TIGER annotation scheme, this label was extended to titles of films, books, exhibitions etc. that have a complex, often sentence-like structure. In these cases, the title is structurally annotated and receives an additional parent-label PN.

Figures 7 and 8 show the different treatment of structured proper nouns. The TIGER annotation scheme thus facilitates the identification of names that do not feature the part-of-speech label NE (proper noun) in one of their terminal nodes.

	precision	recall	F-score
OP	91.14%	71.64%	80.22%
CVC	83.33%	80.00%	81.63%
sec SB	96.51%	91.21%	93.78%
sec OA	100.00%	100.00%	100.00%
sec HD	86.67%	78.79%	82.54%
sec MO	80.00%	77.78%	78.87%
APP	91.27%	90.55%	90.91%
PAR	93.42%	65.74%	77.17%
'new' PN	100.00%	62.50%	76.92%
'old' PN	93.91%	96.98%	95.42%
overall	92.84%	94.97%	93.89%

Table 1: Precision, recall and F-scores for the new labels

4. Assessment of the annotation scheme

4.1. Evaluation

For the evaluation of the changes described in section 3., a sample of 2000 annotated sentences has been extracted. We compared two versions of this sample, namely the first version of the annotation and the final version which was annotated and compared by two annotators. The latter is taken as the true annotation. The first versions were taken from three different annotators in order to level out personal preferences. We computed precision and recall for the first annotations compared to the final version in order to investigate how precise and how comprehensive the annotation can be performed on the basis of the new rules in the annotation scheme. Additionally, we also used the F-score value (harmonic mean of precision and recall), which is an appropriate measure for inter-annotator agreement.

Table 1 contains the precision, recall and F-score values for the new labels concerning verb-subcategorization (OP, CVC), appositions and parentheses (APP, PAR), secondary edge labels (sec SB, sec OA, sec HD, sec MO) as well as proper nouns ('new' PN). For the sake of comparison, we included precision, recall and F-scores for two additional features, namely for regular proper nouns ('old' PN) and for the used corpus sample, taking into account all nodes, edges and labels (overall).

Comparing the values for the new labels, we can see that precision is in most cases higher than recall. This result shows that the annotators do not have problems with assigning the labels correctly in simple cases. Nevertheless, they frequently miss relevant matches, which is reflected in the low recall values. Thus, the new rules guarantee a high degree of correctness for non-ambiguous phenomena (i.e. high precision) whereas the differentiation between labels in unclear cases still poses problems (i.e. low recall).

The difference between precision and recall is not in all cases significant. However, the fact that precision is higher than recall is a striking pattern in table 1. The only exception for the new labels occurs for secondary edge direct objects (sec OA); for this phenomenon, we could not find an explanation.

In contrast to these findings, the two reference values ('old' PN and overall) show higher values for recall than

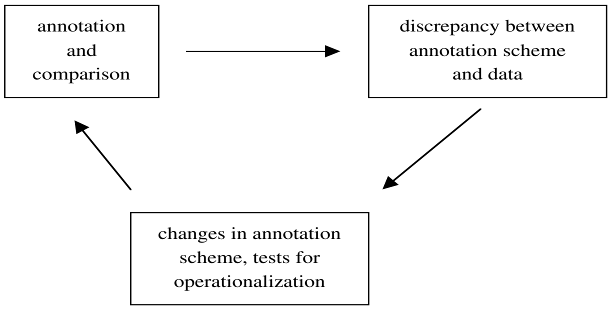

Figure 9: Stages in the development of the annotation scheme

for precision. The reason for this could be that for the older labels that have been used for years, there are not that many unclear phenomena anymore since the annotators developed strategies and rules to disambiguate difficult cases. This assumption is also supported by the higher F-scores for the reference values; with only one exception (the already-mentioned sec OA values), the F-scores for the new labels are considerably lower than the reference F-scores. The reason for this is illustrated in figure 9: there are different stages in the development of the annotation scheme and thus the annotation of the treebank. First, rules are developed and then used in the annotation of the corpus and the subsequent comparison between the annotators. In the course of the annotation and comparison, discrepancies between the annotation scheme and the corpus data become obvious. These discrepancies lead to changes in the scheme and to the development of tests for better operationalization. This cycle results in the cross-fertilization of the corpus and the annotation scheme. Therefore, the scheme improves with the number of times this cycle is run through. Therefore, we expect the new labels to show higher recall values once they have passed through the different stages of this development cycle.

4.2. Problem discussion

In the following, the annotation changes described in section 3. will be critically discussed, taking into account the values that were computed in section 4.1.. We will also show some problems that arise from the new annotation scheme. Furthermore, this section contains ideas on how to improve some of the new rules and on their extension to other phenomena.

4.2.1. Verb-subcategorization

Generally speaking, the annotation concerning the verb-subcategorization still poses some problems. The tests for the labels OP and CVC are apparently not yet sufficiently operationalized in order for the annotators to reach a high level of agreement. The F-score values for the verb-subcategorization labels are among the lowest in table 1.

One problem arises from the use of prepositional objects in NPs. If the head of an NP is derived from a verb, the PP-complement that would be labeled 'OP' in connection with the verb is also labeled 'OP' in the noun phrase:

- träumen [vom Glück]OP → der Traum [vom Glück]OP
- to dream [of happiness]OP → the dream [of happiness]OP

The problem stems from the large amount of composite nouns in German whose main component is one of those derived nouns, but which are not themselves derived from a verb (e.g. Wunschtraum ('wish-dream')). It is not clear whether they should be treated like the derived nouns (i.e. receive the label OP for their complements) or like other 'regular' nouns (i.e. receive the label MNR for their PP-complements).

Furthermore, it must be considered whether the use of the label OP should be extended to PPs in adjective phrases where the adjective is derived from a verb:

- der [vom Glück]OP? träumende Mann
- the [of happiness]OP? dreaming man (the man who dreams of happiness)

This would clearly add to the consistency of the annotation.

4.2.2. Coordination

As was shown above, the introduction of a new layer of annotation, the secondary edges, is a very useful addition to the annotation scheme. It considerably increases the adequacy of the linguistic description. As can be seen in table 1, the secondary edge labels reach the highest F-scores among the new labels (with one exception). This means that the rules concerning the secondary edges turn out quite satisfactory.

There are, however, some inconsistencies when it comes to the inclusion of superordinated sentences in the coordination. On the one hand, there are sentences like

- Steffi hat [geschlafen und geträumt]CVP
- Steffi has [slept and dreamed]CVP

where there would be no secondary edges to indicate the eliptical character in the second part of the coordination. On the other hand, trees like the one in figure 10 can be found in the corpus where the annotator went to great lengths to indicate the elipsis. One aim in the future development of the annotation scheme must be to find a common rule for these cases which also helps to avoid the excessive use of secondary edges like in figure 10.

Furthermore, it might be a good idea to indicate ambiguities in the attachment of modifiers through the use of secondary edges. This was once tried in NEGRA for a short period of time, but was unfortunately discontinued. This part of the annotation would provide valuable information about attachment ambiguities.

4.2.3. Appositions and parentheses

The distinction between appositions and parentheses is fairly clear in most cases. However, a lot of occurences of the label PAR are not recognized as such by the annotators, which is reflected in the extremely low recall value for this feature. The higher F-score for appositions can be explained by the fact that the label APP has been known to

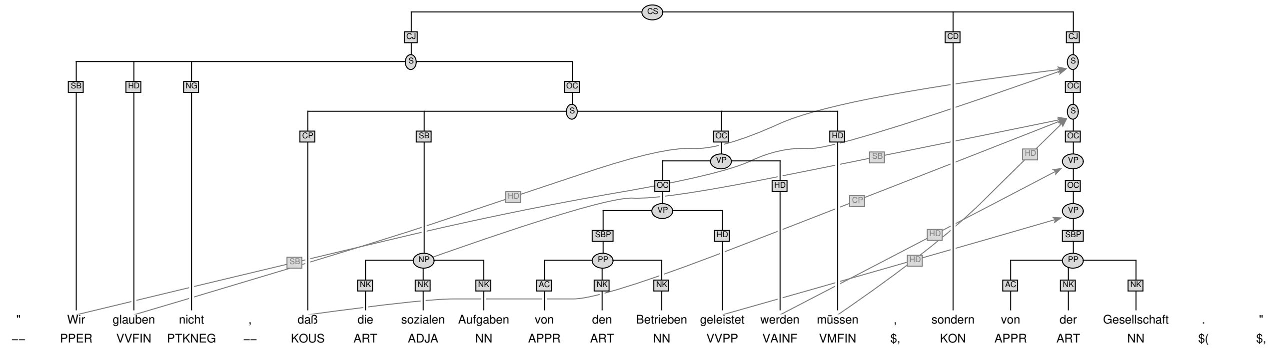

Figure 10: Secondary edges in TIGER

the annotators for a longer time and has therefore passed the development stages displayed in figure 9. Thus, in the annotation of appositions and parentheses, there is still room for improvement concerning the use of the label PAR.

Another weak point is the fact that the label PAR is used in too many cases, which was also the problem with the label APP before: all comments of any kind that do not fit into another category are generally marked with the label PAR. This concerns for instance all kinds of references, addresses, phone numbers etc. that are added in parentheses at the end of a sentence. It would be useful to introduce a rule that limits the usage of the label PAR.

4.2.4. Proper nouns

The treatment of proper nouns still poses some problems for the annotation. Since the recall value for the 'new' proper noun label PN is the lowest in table 1, the rules for the detection of this feature have to be improved.

One of the difficulties is that the rules are not consistent: complex names following terms like 'book', 'film' or 'exhibition' are marked with an additional PN node, but names following the terms 'motto', 'theme' and such are excluded from this rule. The idea behind this is the distinction between a name for a thing and the thing itself, which is too vague and rather a philosophical than a linguistic distinction. Furthermore, it is not yet sufficiently operationalized; it is unclear whether it can be operationalized in a useful way at all.

Another problem concerning proper nouns is the annotation of institutions etc. If the general aim is to mark those names that cannot be identified as such on the part-of-speech level, institutions must in some way be included in this rule. Up to now, the annotation of institution names is highly inconsistent. A phrase like 'Bundesministerium für Arbeit' ('Federal Ministry for Employment') can be found annotated with and without an additional PN node.

In these last cases, it is still unclear whether they are to be treated as 'real' names. One argument in favor of this would be the fact that, once they appear in abbreviated form, they are marked as proper nouns (NE) on the part-of-speech level. Thus, there is no reason why the fully spelt meaning of the abbreviation should not be treated as a proper noun.

5. Summary and outlook

In this paper we presented the TIGER treebank, the largest treebank for German. The sentences are annotated with part-of-speech tags, phrase categories and syntactic functions. It is also possible to encode information about lemmata and morphology. The method of annotation was explained and the different representation formats used in TIGER were presented. The TIGER treebank can be exploited with the TIGERSearch tool, which was also shortly introduced.

We discussed in detail the annotation in TIGER, which is based on the NEGRA annotation scheme. Additionally, we demonstrated the most important differences between the two annotation schemes, which concern verb-subcategorization, the use of secondary edges in coordinations, the differentiation between appositions and parentheses and a more detailed treatment of structured proper nouns.

In an evaluation chapter, precision, recall and F-score were computed for the newly introduced labels and the results were explained. The evaluation showed consistently lower values for recall than for precision. Also, the F-scores for the new features were below the F-score values used as reference. Based on these findings, we discussed some problems that arise from the new annotation rules.

Future work will include the extension of the TIGER treebank to 80.000 sentences altogether and further improvements in the annotation scheme, particularly concerning tests for cases of doubt. It is also planned to introduce further distinctions concerning verbal arguments in order to facilitate the identification of thematic roles (Smith, 2000).

Further information on the corpus, the corpus tools and their availability can be found on the project web page: http://www.coli.uni-sb.de/cl/projects/tiger.

6. References

A. Abeillé, T. Brants, and H. Uszkoreit, editors. 2000a. *Proceedings of the COLING-2000 Post-Conference Workshop on Linguistically Interpreted Corpora LINC-2000*, Luxembourg.

A. Abeillé, L. Clément, and A. Kinyon. 2000b. Building a treebank for french. In *Proceedings of the Second International Conference on Language Resources and Evaluation LREC-2000*, pages 87 – 94, Athens, Greece.

I. Boguslavsky, S. Grigorieva, N. Grigoriev, L. Kreidlin, and N. Frid. 2000. Dependency treebank for russian: Concept, tools, types of information. In *18th International Conference on Computational Linguistics COLING-2000*, Saarbrücken, Germany.

C. Bosco, V. Lombardo, D. Vassallo, and L. Lesmo. 2000. Building a treebank for italian: A data-driven annotation schema. In *Proceedings of the Second International Conference on Language Resources and Evaluation LREC-2000*, pages 99 – 106, Athens, Greece.

T. Brants, R. Hendriks, S. Kramp, B. Krenn, C. Preis, W. Skut, and H. Uszkoreit. 1999. Das NEGRA-Annotationsschema. Technical report, Dept. of Computational Linguistics, Saarland University, Saarbrücken, Germany.

T. Brants. 1997. The NEGRA export format. CLAUS Report 98, Dept. of Computational Linguistics, Saarland University, Saarbrücken, Germany.

T. Brants. 1999. *Tagging and parsing with Cascaded Markov Models - automation of corpus annotation*. Saarbrücken Dissertations in Computational Linguistics and Language Technology (Bd. 6), Saarbrücken, Germany: German Research Center for Artificial Intelligence and Saarland University.

T. Brants. 2000. TnT – a statistical part-of-speech tagger. In *Proceedings of the Sixth Conference on Applied Natural Language Processing ANLP-2000*, Seattle, WA.

S. Dipper, T. Brants, W. Lezius, O. Plaehn, and G. Smith. 2001. The TIGER Treebank. In *Third Workshop on Linguistically Interpreted Corpora LINC-2001*, Leuven, Belgium.

S. Greenbaum, editor. 1996. *Comparing English worldwide: The International Corpus of English*. Clarendon Press, Oxford, UK.

J. Hajic. 1999. Building a syntactically annotated corpus: The Prague Dependency Treebank. In E. Hajicova, editor, *Issues of valency and meaning. Studies in honour of Jarmila Panevova*, Prague, Czech Republic. Charles University Press.

S. Kramp and C. Preis. 2000. Konventionen für die Verwendung des STTS im NEGRA-Korpus. Technical report, Dept. of Computational Linguistics, Saarland University, Saarbrücken, Germany.

G. Leech. 1992. The Lancaster Parsed Corpus. *ICAME Journal*, 16(124).

W. Lezius and E. König. 2000. Towards a search engine for syntactically annotated corpora. In *Proceedings of the Fifth KONVENS Conference*, Ilmenau, Germany.

W. Lezius, H. Biesinger, and C. Gerstenberger. 2002a. Tiger-xml quick reference guide. Technical report, IMS, University of Stuttgart.

W. Lezius, H. Biesinger, and C. Gerstenberger. 2002b. Tigerregistry manual. Technical report, IMS, University of Stuttgart.

W. Lezius, H. Biesinger, and C. Gerstenberger. 2002c. Tigersearch manual. Technical report, IMS, University of Stuttgart.

M. Marcus, G. Kim, M.A. Marcinkiewicz, R. MacIntyre, A. Bies, M. Gerguson, K. Katz, and B. Schasberger. 1994. The Penn Treebank: Annotating predicate argument structure. In *Proceedings of the ARPA Human Language Technology Workshop*, San Francisco, CA. Morgan Kaufman.

A. Moreno, R. Grishman, S. López, F. Sánchez, and S. Sekine. 2000. A treebank of spanish and its application to parsing. In *Proceedings of the Second International Conference on Language Resources and Evaluation LREC-2000*, pages 107 – 112, Athens, Greece.

K. Oflazer, D. Hakkani-Tür, and G. Tür. 1999. Design for a turkish treebank. In *Proceedings of the Workshop on Linguistically Interpreted Corpora LINC-99*, Bergen, Norway.

O. Plaehn and T. Brants. 2000. Annotate - an efficient interactive annotation tool. In *Proceedings of the Sixth Conference on Applied Natural Language Processing ANLP-2000*, Seattle, WA.

G. Sampson. 1995. *English for the computer. The SUSANNE corpus and analytic scheme*. Clarendon Press, Oxford, UK.

A. Schiller, S. Teufel, and C. Stöckert. 1999. Guidelines für das Tagging deutscher Textcorpora mit STTS. Technical report, University of Stuttgart, University of Tübingen.

W. Skut, T. Brants, B. Krenn, and H. Uszkoreit. 1998. A linguistically interpreted corpus of german newspaper text. In *Proceedings of the Conference on Language Resources and Evaluation LREC-98*, pages 705–711, Granade, Spain.

G. Smith and P. Eisenberg. 2000. Kommentare zur Verwendung des STTS im NEGRA-Korpus. Technical report, University of Potsdam.

G. Smith. 2000. Encoding thematic roles via syntactic categories in a german treebank. In *Proceedings of the Workshop on Syntactic Annotation of Electronic Corpora*, Tübingen, Germany.

H. Uszkoreit, T. Brants, and B. Krenn, editors. 1999. *Proceedings of the Workshop on Linguistically Interpreted Corpora LINC-99*, Bergen, Norway.

W. Wahlster, editor. 2000. *Verbmobil: Foundations of Speech-to-Speech Translation*. Springer, Heidelberg, Germany.

H. Zinsmeister, J. Kuhn, and S. Dipper. 2001. From LFG Structures to TIGER Treebank Annotations. In *Third Wokshop on Linguistically Interpreted Corpora (LINC)*, Leuven, Belgium.