The SALSA Corpus: a German corpus resource for lexical semantics

Aljoscha Burchardt, Katrin Erk, Anette Frank*,
Andrea Kowalski, Sebastian Padó, and Manfred Pinkal

Dept. of Computational Linguistics, Saarland University
*Language Technology Lab, DFKI GmbH
Saarbrücken, Germany
{albu, erk, frank, kowalski, pado, pinkal}@coli.uni-sb.de

Abstract

This paper describes the SALSA corpus, a large German corpus manually annotated with role-semantic information, based on the syntactically annotated TIGER newspaper corpus (Brants et al., 2002). The first release, comprising about 20,000 annotated predicate instances (about half the TIGER corpus), is scheduled for mid-2006. In this paper we discuss the frame-semantic annotation framework and its cross-lingual applicability, problems arising from exhaustive annotation, strategies for quality control, and possible applications.

1. Introduction

We describe the SALSA corpus, a German resource which adds word sense and semantic roles to the syntactically analysed TIGER corpus (Brants et al., 2002). The annotation is based on the frame semantics paradigm (Fillmore, 1985). A first release, comprising about 20,000 annotated predicate instances is scheduled for mid-2006.

SALSA addresses a twofold problem of lexical semantics, the scarcity of both annotated corpora and lexica for virtually all languages but English. The availability of treebanks was the driving force behind the recent success of data-driven models in syntax. Likewise, large corpora with rich and reliable semantic annotation are the prerequisite for replicating this success for lexical semantics. Also, the lack of large domain-independent lexica is a major bottleneck for the development of more knowledge-rich NLP applications. Consequently, some of the main uses we envision for the SALSA corpus are the utilisation as a data source for semantics-based NLP applications and the extraction of data for a semantic lexicon.

The semantic annotation of the SALSA corpus can be exploited for corpus-linguistic studies focusing on lexical semantics, syntax-semantics linking properties, or noncompositional phenomena, such as idiomatic or metaphoric expressions, which have been specifically marked in the corpus. Lastly, frame semantics, as a common, largely language-independent word sense and role inventory (Boas, 2005), holds great promise for the cross-lingual analysis and application of lexical semantic information.

Plan of the paper. We first introduce the basics of frame-semantic annotation and describe the concrete annotation scheme used for semantic annotation of the SALSA corpus. Next we discuss the cross-lingual applicability of FrameNet, problems arising from exhaustive annotation, and strategies for quality control. We conclude by listing present and future uses of the corpus.[1]

2. Frame Semantics

Frame Semantics (Fillmore, 1985) seeks to describe the meaning of a sentence as it is actually understood by characterising the background knowledge necessary to understand each expression. It represents this background knowledge in the form of *frames*, conceptual structures modelling prototypical situations. In text, a frame is *evoked* by a word or expression. Table 1 shows the frame STATEMENT, which describes a communication situation and which is evoked by verbs such as *acknowledge* or *admit*, and by nouns such as *affirmation*. Each frame has its own set of semantic roles, called *frame elements* (*FEs*); they are the participants and props in the abstract situation described. In the case of STATEMENT, they are the SPEAKER and ADDRESSEE of the statement, the MESSAGE conveyed and its TOPIC.

Frame: STATEMENT		
This frame contains verbs and nouns that communicate the act of a SPEAKER to address a MESSAGE to some ADDRESSEE using language. A number of the words can be used performatively, such as *declare* and *insist*.		
Frame Elements	SPEAKER	**Evelyn** said she wanted to leave.
	MESSAGE	Evelyn announced **that she wanted to leave**.
	ADDRESSEE	Evelyn spoke **to me** about her past.
	TOPIC	Evelyn's statement **about her past**
	MEDIUM	Evelyn preached to me **over the phone**.
Predicates	acknowledge.v, acknowledgement.n, add.v, address.v, admission.n, admit.v, affirm.v, affirmation.n, allegation.n, allege.v, announce.v, announcement.n, assert.v, assertion.n, attest.v, aver.v, avow.v, avowal.n, ...	

Table 1: Example frame from the FrameNet database

The Berkeley FrameNet project (Baker et al., 1998) is developing a frame-semantic lexicon for the core vocabulary of English. The current release contains some 600 frames and 8,700 lexical units. FrameNet organises frames in a hierarchy which also provides a role mapping between linked frames. Frames are illustrated with annotated examples from the British National Corpus; FrameNet currently contains about 133,000 annotated sentences.

The annotation of predicate-argument structure in general, and of FrameNet in particular, is interesting for its intermediate position between syntax and "deep" semantics. It generalises across near meaning-preserving transformations such as verbal alternations (examples (1) and (2)) or nominalisations (example (3)). The semantic role labels characterise the relationship between predicate and argu-

[1] We acknowledge the funding of the German Science Foundation DFG for the SALSA project (Title PI 154/9-2).

ment as well as relationships among arguments. This provides a handle on modelling inferences about role-fillers: for example, the THEME of a GIVING event is the object that changes possessors. In addition, FrameNet's hierarchy of frames, which is mirrored in a hierarchy of roles, enables further generalisation over roles and inferences over frames. At the same time, frame semantics disregards problems of deep semantic analysis such as modality, negation, or scope ambiguity.

(1) [Peter]$_{\text{AGENT}}$ hit$_{\text{CAUSE_IMPACT}}$ [the ball]$_{\text{IMPACTEE}}$.
[The ball]$_{\text{IMPACTEE}}$ was hit$_{\text{CAUSE_IMPACT}}$.

(2) [Sue]$_{\text{DONOR}}$ gave$_{\text{GIVING}}$ [Mary]$_{\text{RECIPIENT}}$ [a book]$_{\text{THEME}}$.
[Sue]$_{\text{DONOR}}$ gave$_{\text{GIVING}}$ [a book]$_{\text{THEME}}$ [to Mary]$_{\text{RECIPIENT}}$.

(3) [Evelyn]$_{\text{SPEAKER}}$ spoke$_{\text{STATEMENT}}$ [about her past]$_{\text{TOPIC}}$.
[Evelyn's]$_{\text{SPEAKER}}$ statement$_{\text{STATEMENT}}$ [about her past]$_{\text{TOPIC}}$

(4) [The minister]$_{\text{EMPLOYEE}}$ was dismissed$_{\text{FIRING}}$.
[The minister]$_{\text{EMPLOYEE}}$ was given the boot$_{\text{FIRING}}$.

FrameNet's frame elements are local to individual frames. This avoids the commitment to a small set of universal roles, whose specification has turned out to be controversial in the past. The coarse-grained frame descriptions generalise over lexical variation (example (4)). In contrast, PropBank (Palmer et al., 2005) focuses on the mapping of syntax to semantic roles for individual predicates, not committing itself to higher-level generalisations. As a consequence, PropBank roles have only a *verb-specific* semantic interpretation (cf. Ellsworth et al. (2004)).
A central point of interest is the multilingual dimension of FrameNet. A number of projects are investigating the use of FrameNet frames of languages other than English, such as German (SALSA), Spanish (Subirats and Petruck, 2003) and Japanese (Ohara et al., 2004). Even though FrameNet frames have turned out to be to a large extent language-independent, they are not fully interlingual. In fact, frame-semantic analyses of sentence translation pairs allow the investigation of similarities and differences in how different languages express similar meaning (Padó and Erk, 2005).

3. SALSA: Frame-semantic corpus annotation for German

The aim of the SALSA project is to create a German lexical semantics resource, following the theoretical framework of frame semantics. Similar to PropBank (Palmer et al., 2005), SALSA has chosen a corpus-based approach, and extends an existing German treebank, the TIGER treebank (Brants et al., 2002), with lexical semantic annotations.

Annotation scheme. We annotate frame-semantic information on top of the syntactic structure of the TIGER corpus, with a single flat tree for each frame. The root node is labelled by the frame name, the edges are labelled with the names of the frame elements. The FEs refer to syntactic constituents. Figure 1 shows a simple annotation instance: the verb *antwortet* ("answers") introduces the frame COMMUNICATION_RESPONSE. The NP subject *die Branche* is

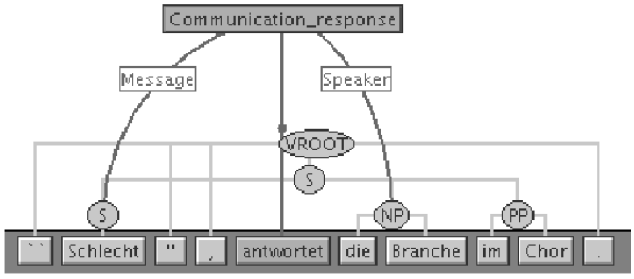

Figure 1: Annotation example: "'Badly', the industry sector answers in unison."

annotated as SPEAKER and *schlecht*, under an S node, as MESSAGE. The picture is a screenshot of SALTO, a graphical annotation tool (Burchardt et al., 2006), which stores annotations in SALSA/TIGER XML (Erk and Padó, 2004).

Annotation process. Annotation proceeds one predicate at a time and is *exhaustive* in that all instances of a predicate are annotated. Each predicate is annotated by two independent annotators. In a subsequent *adjudication* step, cases of disagreement between annotators are resolved manually.

Release coverage. SALSA annotates primarily predicates with a proper predicate-argument structure. For the time being, we concentrate on verbal predicates, parallel to PropBank practice. Our first release, which is scheduled for mid-2006, will consist of about 500 German verbal predicates of all frequency bands plus some deverbal nouns, with a total size of around 20,000 annotated instances.

SALSA and FrameNet. FrameNet's aim is to create a linguistically structured network of frames and roles by exploring semantic fields and their linguistic realisation possibilities. SALSA builds on the FrameNet inventory frames and roles, but follows a *corpus-based* approach.
The fact that our German corpus annotation is based on frames and roles that were created for English raises the question of the applicability of frame semantic descriptions to other languages, i.e. the multilingual dimension of frame semantics in general, and the FrameNet resource in particular. Moreover, applying the (still incomplete) FrameNet lexicon presents us with the challenge of gaps in the inventory of frames, and the problem of "grey areas" and productive usages usually not described in a lexicon. Both of these aspects will be discussed in detail below.

4. How well does FrameNet work for exhaustive annotation of German?

In our experience, the vast majority of FrameNet frames can be used fortuitously to describe German predicate-argument structure. Nevertheless, some FrameNet frames required adaptation for SALSA annotation (Sec. 4.1), and we had to introduce new frames (Sec. 4.2).

4.1. Cross-lingual divergences

In applying FrameNet frames to German data, we found three main problem sources.

Ontological distinctions between frame elements. In some cases, FrameNet uses ontological criteria to differentiate between closely related, mutually exclusive FEs. For example, consider the frame ASSISTANCE:

A HELPER helps a CO_AGENT to complete a GOAL that the CO_AGENT has, by participating in some action with the CO_AGENT. A FOCAL_ENTITY that is involved in reaching the GOAL may stand in for it.

The underlined FEs are metonymically related and can be seen as instances of one more abstract role. In English, the distinction can be made on syntactic grounds, since the FOCAL_ENTITY is usually a *with*-PP, while the GOAL is a VP or S. In German, however, *helfen (to help)* frequently occurs with a *bei*-PP containing a deverbal noun:

(5) Luise hilft Hans, [das Geschirr zu spülen]$_{\text{GOAL}}$.

(6) Luise hilft Hans [mit dem Geschirr]$_{\text{FOCAL\_ENTITY}}$.

(7) Luise hilft Hans [beim Geschirrspülen]$_{??}$.

The role in example (7) is neither a GOAL nor a FOCAL_ENTITY, but an action in which the HELPER participates and which is suitable for attaining the GOAL. In the case at hand, we defined a frame element which generalises over GOAL, FOCAL_ENTITY as well as actions as in (7). We proceed in a similar fashion for related cases which involve purely ontological role distinctions.

Missing frame elements. The use of dative objects is much less restricted in German than in English. This leads to problems, when a frame fits a sense of a German predicate, but lacks the frame element that can be realised as a dative in German. An example is the frame TAKING, in which an AGENT takes possession of a THEME by removing it from a SOURCE. In English, the SOURCE, usually realised as a *from*-PP, can be either a source location or a former possessor; both together can be expressed only clumsily. In contrast, the German verb *nehmen* can realise location and possessor simultaneously:

(8) Er nahm [ihm]$_{\text{POSS.}}$ [das Bier]$_{\text{TH.}}$ [aus der Hand]$_{\text{SRC.}}$
He took him the beer out of the hand

To handle such cases, we add new roles – here a POSSESSOR role, thereby splitting the FrameNet SOURCE role into a location-type SOURCE and a distinct POSSESSOR.

Differences in lexical realization patterns. At times, German verbs show patterns which run counter to the frame distinctions made on English data. An example is the German *fahren*, which encompasses both English *drive* (frame OPERATE_VEHICLE, as the driver) and *ride* (frame RIDE_VEHICLE, as a passenger). In German, the context usually does not disambiguate between the two frames, which makes it impossible to make the decision reliably. In the case at hand, FrameNet has introduced the frame USE_VEHICLE, which subsumes both OPERATE_VEHICLE and RIDE_VEHICLE. While the frame is unlexicalised for English, it is the right level to describe the meaning of German *fahren*. In general, such cases need to be discussed from a multilingual perspective; in the ongoing annotation, we resort to underspecification (see Sec. 5) for such cases.

4.2. Extending FrameNet coverage

Recall that SALSA annotation proceeds one predicate at a time (Sec. 3). Since FrameNet does not yet cover the complete "word sense space", we have to check for each new predicate whether all senses are covered. To this effect, we inspect a number of TIGER occurrences of each new predicate before actual annotation begins. We found that a sample size of twenty is a reasonable compromise between keeping the effort practicable and encountering the most important senses.

Frame: RECHNEN.UNKNOWN3		
An Item is construed as an example or member of a specific category. In contrast to Categorisation, no Cognizer is involved. In contrast to Membership, the Category does not have to be a social organisation.		
FEs	ITEM	**Die Philippinen und Chile** rechnen zu den armen Ländern der Region.
	CATEGORY	Die Philippinen und Chile rechnen **zu den armen Ländern der Region.**

Table 2: Example of a proto-frame. "The Philippines and Chile are counted among the region's poor countries."

For each instance, we check whether some FrameNet frame applies. The decision is based on the criteria detailed in Ellsworth et al. (2004): Does the meaning of the instance meet the frame definition? Can all important semantic arguments of the instance be described in terms of the frame elements? In cases of doubt, we also check annotated FrameNet example sentences for similar usages.

We group instances with non-covered readings into "sense groups" and construct a *predicate-specific proto-frame* for each group. Figure 2 shows a proto-frame we constructed for the *to be counted (among a group)* sense of *rechnen*. Similar to FrameNet frames, the SALSA proto frames have a textual definition, a set of roles with FrameNet-style names, and annotated example sentences. The proto-frames follow a naming convention, e.g. RECHNEN.UNKNOWN3 for the third such frame for the predicate *rechnen*.

Although SALSA is not a lexicographic project, our predicate-specific proto-frames can provide input for the further development of FrameNet: We attempt to keep proto-frames at roughly the same level of granularity as FrameNet frames. In addition, we list frame-to-frame relations for proto-frames to indicate their relationship to both FrameNet frames and other proto-frames. E.g., for RECHNEN.UNKNOWN3 we record that it is identical to a proto-frame for *zählen*; in the example sentence, *rechnen* can be paraphrased by *zählen*.

Statistics. For a dataset of 476 German predicates, for which annotation was finalised by the time of writing, we counted 18,500 instances with 628 different frames. 252 were FrameNet frames, and 373 new proto-frames. The average number of frames per predicate was 2.8, composed of 2.0 FrameNet frames and 0.8 proto-frames. In other words, somewhat less than one third of the predicate senses in our corpus was not covered by FrameNet. Not surprisingly, the actual number of senses varies greatly between individual predicates; the lemma with the highest number of frames is *kommen*, with 39 frames (29 FrameNet + 10 proto).

The average polysemy in SALSA (2.8) is higher than the current average WordNet verb polysemy (2.2); this is at least partly due to our treatment of idioms and metaphoric readings as additional senses of predicates.

	246 Lemmas		*nehmen*	
	Number	%	Number	%
Compositional	**4638**	**85.7**	**42**	**17.4**
Metaphor	369	6.8	38	15.8
Support	326	6.0	132	45.8
Idiom	79	1.5	29	12.0
LC	**774**	**14.3**	**199**	**82.6**
Total	**5412**	**100.0**	**241**	**100.0**

Table 3: Phenomena with limited compositionality (LC)

5. Varying degrees of compositionality in exhaustive annotation

In standard annotation cases, there is a strong parallelism between syntactic and semantic structure: a single-word predicate lexically introduces a frame, whose frame elements are syntactic arguments (i.e. subcategorised for). Figure 1 shows an example of such a case. However, due to our exhaustive annotation policy, we frequently encounter cases of *limited compositionality* in which frame choice, argument choice, or both, diverge from this simple picture. The main phenomena are support verb constructions, idioms, and metaphors. Their frequencies in a corpus sample of around 5,400 instances are shown in Table 3. Almost one seventh of this sample constituted instances of these phenomena. For high-frequency, and therefore highly polysemous, verbs such as *nehmen (to take)*, these phenomena even constitute the majority of instances.

Support Verb Constructions. A support verb constructions (SVC) is a combination of a verb with a "bleached" or abstract meaning (e.g. causation or perspectivisation) with a predicative noun, typically its object, which constitutes the semantic head of the phrase, and should be treated as frame-evoking element. An example is *Abschied nehmen (to take leave)*. Often, the SVC can be paraphrased with a morphologically related verb (*sich verabschieden*). Currently, SALSA annotates the verbal parts of SVCs with a pseudo frame SUPPORT, whose only FE, SUPPORTED, points to the noun. This annotation makes SVCs retrievable and thus available for a later, more elaborate analysis of the syntax-semantics interaction between verb and noun.

Idioms. We use three criteria for identifying idioms: Idioms are multi-word expressions which are (a) (for the most part) fixed, (b) introduce the meaning as a whole, and (c) whose understood meaning is not synchronically recoverable from their literal meaning. An example is *Nachteile in Kauf nehmen*, literally *to take disadvantages into purchase*, meaning *to put up with disadvantages*. Our annotation scheme for idioms is to annotate the complete multiword expression as the frame-evoking element; arguments do not require special treatment.

Metaphors. Metaphors are distinguished from idioms by the existence of a figurative reading which is recoverable from their literal meaning. This characterisation, which corresponds well to Lakoff's ideas on metaphorical transfer involving source and target domains (Lakoff and Johnson, 1980), suggests the simultaneous annotation of two frames: a *source* frame to represent the literal meaning, and a *target* frame to represent the figurative meaning. As an example, consider *unter die Lupe nehmen (to put* (literally: *take*) *under a magnifying glass)*. The source frame is TAKING, and the target frame is SCRUTINY, which models the construction of this metaphor as a transfer from a (concrete) putting event to a (more abstract) investigation event.

We attempt to annotate both frames for all metaphorical instances, and mark their status with frame flags *Source* and *Target*. Being the result of a complex interpretation process, the target meaning is often difficult to describe. We annotate these cases with the source frame only to sustain annotation speed. In a later stage, these samples can be retrieved for a more comprehensive analysis.

Transfer Schemes for Metaphors. Source and target frames describe complementary properties of metaphors: The source frame models the syntactic realization patterns of arguments, while the target frame captures the understood meaning. Those instances which have received source *and* target frames can be used to study *transfer schemes*, including information about *argument change*. The SALSA annotation seems well-suited for this task, since frames as sense classes provide an empirically founded, fine-grained vocabulary to describe transfer processes.

In simple cases, the transfer establishes a direct correspondence between source and target frames, including all arguments. In the example *Das Postfach explodiert (The mailbox explodes)*, the source frame CHANGE_OF_PHASE with its role UNDERGOER directly maps onto the target frame EXPANSION with the role ITEM. As a more complex case, consider *unter eine* **starke** *Lupe nehmen (to put under a* **strong** *magnifying glass)*. The corresponding transfer scheme in Fig. 2 shows a case of *argument incorporation*: the GOAL role of PLACING is absorbed in the frame-evoking element of SCRUTINY; in addition, the modifier *starke (strong)*, which does not fill a role on the source side, becomes the DEGREE in the target frame.

Transfer schemes such as the one shown here do not answer the question as to which factors trigger the metaphorical transfer for a specific utterance. However, they can model the interpretation process of metaphors to a certain degree, and provide a description of the relation between source and target for specific metaphors, which makes it possible to express generalisations over patterns of role shift.

Vagueness. It is a well-known fact that in semantic annotation there are cases of vagueness in which the assignment of only a single label to a markable would not be appropriate (Kilgarriff and Rosenzweig, 2000). For such cases, annotators should be able to assign more than one label. This makes it possible to retrieve vague cases, and it avoids forcing them to make impossible choices.

SALSA annotation faces the problem of vagueness both at the level of frames and frame elements. As an example for frames, *die Tür zuschlagen* (*slam the door*) has aspects of CAUSE_IMPACT (the door is caused to slam into its frame) but also of CLOSURE (the door is being closed). As an example for frame elements, consider the metonymic sentence (9): *the motion* describes the MEDIUM used to convey the demand, but metonymically it also refers to the SPEAKER.

(9) Die nachhaltigste Korrektur fordert **ein Antrag**
The most radical change is demanded by **a motion**

In cases like these, SALSA annotators assign more than one

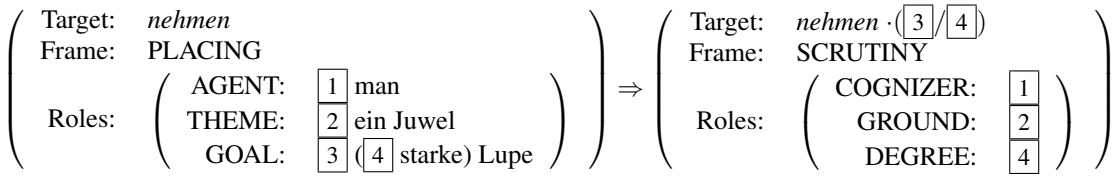

Figure 2: Transfer scheme for *Die Klangkultur ist ein Juwel, das man getrost unter eine starke Lupe nehmen kann.* ("Their sound is a jewel which stands up to any scrutiny.")

frame (or more than one frame element), connecting the multiple assignments by an *underspecification* link. Underspecification does not have an a priori disjunctive or conjunctive interpretation, since it has been argued (Kilgarriff and Rosenzweig, 2000) that it is impossible for annotators to decide reliably between the two.

Underspecification is particularly useful to represent borderline instances of phenomena with limited compositionality. Notorious cases are the distinction between support constructions and metaphors, and between (transparent) metaphors and (no longer transparent) idioms.

6. Quality Control

The four-eye principle. SALSA aims at guaranteeing quality by double, independent analysis of all data. Each dataset for a given predicate is annotated independently by two annotators in changing pairs. Through this *double annotation*, a fair number of annotation mistakes can be detected automatically. After annotation, the two versions of a dataset are merged into a single copy in which annotation differences are marked. Next, the differences in the merged version are subject to *double adjudication*, in which conflicts are resolved manually. Since frame-semantic annotation is a novel task that combines word sense and structural annotation, this allows us to assess its difficulty. Differences remaining after adjudication are resolved jointly in a final *meta-adjudication* step.

Computing Agreement. On the basis of two independently annotated and two adjudicated versions, we compute *inter-annotator agreement* and *inter-adjudicator agreement.* [2] We consider frame selection and role labelling individually, due to their different characteristics.

Our inter-annotator agreement is 85% on frames and 86% on roles. Inter-adjudicator agreement is 97% on frames and 96% on roles. In other words, annotators agree on more than 4/5ths of all instances; adjudication creates consensus for another 4/5ths of the disagreements.

[2] It is best practice for annotation projects to report chance-corrected agreement; the most widely measure is the kappa statistic (Siegel and Castellan, 1988). However, kappa assumes a very restricted annotation process, in which a single label is chosen from a globally fixed pool for each annotated instance. This assumption is appropriate for neither of our two subtasks: In frame selection, annotators can use underspecification to dynamically create new classes, or skip instances requiring additional proto-frames; in role labelling, the majority of instances (syntactic constituents) receives the "no role" label, which leads to misleadingly high agreement numbers. This problem is pertinent to to many recent, more complex annotation efforts, such as the Penn Discourse Treebank (Miltsakaki et al., 2004). Following their argumentation, we do not report kappa, but percentage agreement according to a strict evaluation metric (labelled exact match).

Remaining disagreements. Almost all disagreements which remain after adjudication are truly difficult cases. Many are *idiosyncractic problems*, i.e. problems with particular instances. Examples are referential ambiguities, which can lead to ambiguous role assignments. A second category consists of *conceptual problems* with respect to the FrameNet inventory. Examples are systematic problems in distinguishing roles (Sec. 4.1), or usages which meet frame descriptions only partially, or else combine aspects of several frames (Sec. 5). For some cases, underspecification can be used as a "last resort" to represent at least the uncertainty about the correct analysis.

Dynamics of underspecification. While underspecification is a well-motivated device (see Sec. 5 and the last paragraph), there is the danger that annotators use it as a "blanket annotation" for cases in which they feel subjectively uncertain. By tracing the frequency of underspecified annotation across adjudication and meta-adjudication, we can track whether underspecification in annotation is confirmed, i.e. a valid expression of vague or complex meanings. We found that frame underspecification is about three times as frequent as frame element underspecification; also, the amount of frame underspecification increases slightly in adjudication, while half of the frame element underspecification is rejected. However, there was a high fluctuation across individual predicates.

This confirms our intuition that frame choice is the more difficult problem, and justifies the use of underspecification. Especially the phenomena of Sec. 5 contribute to the complexity of frame choice. The low number of underspecified frame element annotations, and their decrease in adjudication, suggests that problems in frame element annotation are less fundamental. Finally, the difficulty of annotation is highly lexicalised, varying across predicates.

Limits of double analysis. Quality control using inter-annotator agreement can only identify errors caused by individual annotation differences between annotators. If both annotators make the same error, it cannot be detected. This limits the effectiveness of quality control by inter-annotator agreement with regard to systematic mistakes.

For this reason, we draw random samples for all completely annotated predicate-frame-pairs, which are inspected for possible systematic annotation mistakes. We have also experimented with *intra*-annotator agreement, trying to detect errors by finding "outliers" with non-uniform behaviour. However, due to the highly lexicalised nature of semantic annotation, even correctly annotated datasets can show non-uniformities, which leads to false positives.

A currently unsolved problem is how consistency can be guaranteed across different predicates annotated with the same FrameNet frame, especially in the face of difficult distinctions, e.g. between frame elements.

7. Using the Corpus

Querying the corpus. The first SALSA release will be distributed in SALSA/TIGER XML format (Erk and Padó, 2004), free for academic research. The XML format can be accessed using script languages with XPATH functionality, such as XSLT. Through a transformation of SALSA/TIGER XML to Berkeley FrameNet XML, the corpus can also be comfortably queried using the web-based FrameSQL query tool (Sato, 2003) which allows cross-lingual contrastive browsing of semantic valencies.

Corpus-based studies. The corpus offers frame semantic annotations including a variety of special phenomena (supports, metaphors, etc.). This allows the user to conduct corpus-based studies focusing on semantic structures alone, or their syntax-semantics linking patterns.

Lexicon. Generalisations over semantic structures and their linking properties as encoded in the corpus can, more generally, be represented in the form of a lexicon. SALSA is currently designing a German frame-based lexicon model in a description logic framework. This model will include frame descriptions, their syntax-semantics linking patterns with frequency distributions, as well as further information, such as selectional preferences. The lexicon descriptions will be extracted from the corpus annotations and at the same time will provide back-references to the annotation instances, thus "grounding" the lexicon in the corpus.

Applications using frame semantics. A well-known use for corpora with role-semantic annotation is the training of shallow semantic parsers (Gildea and Jurafsky, 2002; Erk and Padó, 2006). Frame-semantic annotation has also been used as a flat semantic projection layer on top of a symbolic LFG grammar (Frank and Erk, 2004; Frank and Semecky, 2004), with interfaces to the WordNet and SUMO ontologies (Burchardt et al., 2005a).
In the context of the recent RTE challenge, frame-semantic representations have been applied successfully to approximate textual entailment (Tatu and Moldovan, 2005; Burchardt and Frank, 2006). Frame-based processing has further been applied both for textual QA (Fliedner, 2006) and as a QA interface to structured knowledge bases (Frank et al., 2006). Further research directions include the study of interactions of frame structures with discourse phenomena (Burchardt et al., 2005b).

8. References

C. F. Baker, C. J. Fillmore, and J. B. Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of COLING-ACL 1998*, Montréal, Canada.

H. C. Boas. 2005. Semantic frames as interlingual representations for multilingual lexical databases. *International Journal of Lexicography*, 18(4):445–478.

S. Brants, S. Dipper, S. Hansen, W. Lezius, and G. Smith. 2002. The TIGER treebank. In *Proceedings of the Workshop on Treebanks and Linguistic Theories*, Sozopol.

A. Burchardt and A. Frank. 2006. Approaching textual entailment with LFG and FrameNet frames. In *Proceedings of RTE-2 Workshop*, Venice, Italy.

A. Burchardt, K. Erk, and A. Frank. 2005a. A WordNet Detour to FrameNet. In *Sprachtechnologie, mobile Kommunikation und linguistische Resourcen*, volume 8 of *Computer Studies in Language and Speech*. Peter Lang, Frankfurt.

A. Burchardt, A. Frank, and M. Pinkal. 2005b. Building text meaning representations from contextually related frames – a case study. In *Proceedings of IWCS 2005*, Tilburg.

A. Burchardt, A. Frank, K. Erk, A. Kowalski, and S. Padó. 2006. SALTO – a versatile multi-level annotation tool. In *Proceedings of LREC 2006*, Genoa, Italy.

M. Ellsworth, K. Erk, P. Kingsbury, and S. Padó. 2004. PropBank, SALSA and FrameNet: How design determines product. In *Proceedings of the Workshop on Building Lexical Resources From Semantically Annotated Corpora at LREC 2004*.

K. Erk and S. Padó. 2004. A powerful and versatile XML format for representing role-semantic annotation. In *Proceedings of LREC 2004*, Lisbon, Portugal.

K. Erk and S. Padó. 2006. Shalmaneser – a toolchain for shallow semantic parsing. In *Proceedings of LREC 2006*, Genoa, Italy.

C. J. Fillmore, C. R. Johnson, and M. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography*, 16:235–250.

C. J. Fillmore. 1985. Frames and the semantics of understanding. *Quaderni di Semantica*, IV(2).

G. Fliedner. 2006. Towards natural interactive question answering. In *Proceedings of LREC 2006*, Genoa, Italy.

A. Frank and K. Erk. 2004. Towards an LFG syntax-semantics interface for frame semantics annotation. In *Proceedings of CICLing 2004*.

A. Frank and J. Semecky. 2004. Corpus-based Induction of an LFG Syntax-Semantics Interface for F rame Semantic Processing. In *Proceedings of LINC 2004*, Geneva, Switzerland.

A. Frank, H.-U. Krieger, F. Xu, H. Uszkoreit, B. Crysmann, B. Jörg, and U. Schäfer. 2006. Question answering from structured knowledge sources. *Journal of Applied Logic, Special Issue on Questions and Answers: Theoretical and Applied Perspectives*. to appear.

D. Gildea and D. Jurafsky. 2002. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3):245–288.

A. Kilgarriff and J. Rosenzweig. 2000. Framework and results for English Senseval. *Computers and the Humanities*, 34(1-2).

G. Lakoff and M. Johnson. 1980. *Metaphors we live by*. University of Chicago Press, Chicago, IL.

E. Miltsakaki, R. Prasad, A. Joshi, and B. Webber. 2004. Annotating discouse connectives and their arguments. In *Proceedings of the Workshop on frontiers in corpus annotation at HLT/NAACL*.

K. Ohara, S. Fujii, T. Ohori, R. Suzuki, H. Saito, and S. Ishizaki. 2004. The Japanese FrameNet project: An introduction. In *Proceedings of the Workshop on Building Lexical Resources from Semantically Annotated Corpora at LREC 2004*.

S. Padó and K. Erk. 2005. To cause or not to cause: Cross-lingual semantic matching for paraphrase modelling. In *Proceedings of the Workshop on Cross-Linguistic Knowledge Induction at EUROLAN 2005*.

M. Palmer, D. Gildea, and P. Kingsbury. 2005. The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1).

H. Sato. 2003. FrameSQL: A software tool for the FrameNet database. In *Proceedings of the 3rd Conference of the Asian Association for Lexicography*, Tokyo, Japan.

S. Siegel and N.J. Castellan. 1988. *Nonparametric statistics for the Behavorial Sciences*. McGraw-Hill, 2nd edition.

C. Subirats and M. Petruck. 2003. Surprise! Spanish FrameNet! In *Proceedings of the Workshop on Frame Semantics at the XVII. International Congress of Linguists*, Prague.

M. Tatu and D. Moldovan. 2005. A semantic approach to recognizing textual entailment. In *Proceedings of HLT/EMNLP 2005*, pages 371–378, Vancouver, BC.