PropBank, SALSA, and FrameNet: How Design Determines Product

Michael Ellsworth*, Katrin Erk†, Paul Kingsbury‡, Sebastian Padó†

*International Computer Science Institute
1947 Center St., Berkeley, CA, USA
infinity@icsi.berkeley.edu

†Computational Linguistics, Saarland University
Im Stadtwald 17, 66123 Saarbrücken, Germany
{erk,pado}@coli.uni-sb.de

‡Institute for Research in Cognitive Science
3401 Walnut St., Philadelphia, PA, USA
kingsbur@unagi.cis.upenn.edu

Abstract

We compare three projects that annotate semantic roles: PropBank, FrameNet, and SALSA. The first part of our analysis is a comparison of the different word sense distinction criteria underlying the annotation. Then, we study the effects of these criteria at the level of actual phenomena that require annotation. In particular, we discuss metaphor, support constructions, words with multiple meaning aspects, phrases realizing more than one semantic role, and nonlocal semantic roles.

1. Introduction

During the last few years, corpora with semantic role annotation have received much attention, since they offer rich data both for empirical investigations in lexical semantics and large-scale lexical acquisition for NLP applications.

However, semantic role annotation of text is a complicated endeavor, whose product is deeply influenced by the initial design philosophies and policy choices of a project. We examine key differences between three annotation projects, FrameNet (Johnson et al., 2002), PropBank (Kingsbury and Palmer, 2002), and SALSA (Erk et al., 2003), and the consequences of these differences. After introducing the goals of the projects, we compare the criteria for determining the words senses underlying the annotation. Then, we discuss the consequences of these choices at the level of actual annotation.

2. PropBank, FrameNet and SALSA

FrameNet is primarily a lexicographical project. Its starting point is the observation that words can be grouped into semantic classes, the so-called 'frames', representations for prototypical situations or states. Each frame provides its set of semantic roles. The Berkeley FrameNet project is building a dictionary which links frames to the words and expressions that can introduce them in text. Examples from the BNC (Burnard, 1995) serve to illustrate typical usages.

The more practical aim of PropBank, on the other hand, was to obtain a complete semantic role annotation of the Penn Treebank (Marcus et al., 1994). The PropBank lexicon was added first to facilitate annotation, and later evolved into a resource on its own. No higher-order organization was established at first, so for each unique verb sense, a 'frameset' was constructed that consists of the set of semantic roles and the accompanying syntactic realizations of each.

SALSA uses the FrameNet dictionary as the basis for its annotation but, like PropBank, pursues an exhaustive annotation of its corpus, the TIGER corpus (Brants et al., 2002), a German newspaper corpus. Different from FrameNet, however, SALSA is not committed to always assigning a single sense (frame) to a target expression, or a single semantic role to a constituent. In cases of systematic as well as idiosyncratic ambiguity and vagueness, annotators may assign more than one frame or semantic role and mark the occurrence as being 'underspecified'.

3. Criteria for frameset and frame creation

In this section, we describe the criteria used for grouping instances of role-introducing expressions (targets) into senses, i.e. frames (in FrameNet) and framesets (in PropBank), respectively. SALSA uses FrameNet's criteria.

3.1. PropBank

Since the purpose of the PropBank lexicon was primarily to provide a description of every verb in the Penn Treebank II corpus in all their attested usages, it was kept as agnostic as possible with respect to higher-level generalizations. Recall from above that framesets are verb-specific, and even though polysemous verbs could possess multiple framesets, in general senses were merged into single framesets whenever possible. Distinction of senses, and therefore creation of distinct framesets, was triggered by both syntactic and semantic properties.

One important criterion is the number of possible semantic roles. For example, the verb *afford* is given two framesets, on the basis of pairs of sentences such as:

(1) These days Nissan can **afford** that strategy, even though profits aren't exactly robust. (wsj_0286)

(2) Last year the public was **afforded** a preview of Ms. Bartlett's creation in a tablemodel version, at a BPC

exhibition. (wsj_0984)

Although each sentence has two realized semantic roles, the passive morphology in (2) indicates that a third role is possible. The same is not true for (1), which leads to the creation of two framesets:

afford.01	'be able to sustain the cost of something'
arg0:	entity sustaining cost
arg1:	costly thing

afford.02	'provide, make available'
arg0:	provider
arg1:	thing provided
arg2:	recipient

This pair of sentences also serves to illustrate how it is not the number of roles appearing in any sentence which determines the framing, but the number of possible roles across a variety of syntactic alternations such as active/passive or causative/inchoative.

Even if the number of roles is the same, framesets are also distinguished when the meanings of the usages are sufficiently different, that is, if a relatively proficient speaker of English will be able to distinguish between these senses. For example, the verb *stem* also takes two framesets[1], each with two roles, given pairs of sentences such as:

(3) Travelers Corp.'s third-quarter net income rose 11%, even though claims **stemming** from Hurricane Hugo reduced results $40 million. (wsj_0144)

(4) If the company can start to ship during this quarter, it could **stem** some, if not all of the red ink, he said. (wsj_1973)

PropBank therefore assumes the following two framesets:

stem.01	'arise'
arg1:	entity arising, coming about
arg2:	arising from what?

stem.02	'stanch, cause to stop flowing'
arg0:	causer of non-flowing
arg1:	thing no longer flowing

Because roles are defined per verb, the classification of individual verbs into higher-level classes is not trivial. Most framesets make reference to VerbNet (Kipper et al., 2002) classes, a refinement of Levin's (1993) scheme, and efforts are underway to discover natural classes of verbs based on patterns of usage (Kingsbury and Kipper, 2003).

3.2. FrameNet

FrameNet practice differs fundamentally from the process described for PropBank in not considering syntactic differences (except inasmuch as these correlate with semantics). This means that FrameNet can consider verbs, adjectives, nouns, prepositions, adverbs, and even multiword expressions (such as *pull the wool over X's eyes* in the Prevarication frame) on the same footing, despite any structural differences between them, since it is only their meaning which matters.

FrameNet's semantic criteria for creating frames also differ from those of PropBank in taking the senses as less predefined. FrameNet first collects and analyzes the corpus attestations of target words (or idiomatic phrases) judged to have semantic overlap (as determined by consulting thesauri, dictionaries, and native intuitions). The attestations are divided into semantic groups, noting especially the semantic roles (frame element) of each (but ignoring pragmatic and general constructional differences, as PropBank does), and then combining these small groups into frames. Note that the resulting groupings need not correspond to the initial groupings given by thesauri, etc. The factors which may serve to differentiate or relate the groups of attestations include the following:

1. For the target:
   (a) The basic denotation of the targets may differ, such as in the case of *blue* and *broken* which refer to completely different kinds of states. Obviously this is a diagnostic which is easy in some cases and hard in others. A more difficult case is the basic meanings of *take* vs. *receive*, which share lots of implications about a Theme changing hands. It is simply unclear whether these are exactly the same kind of thing. The difficulty of forming an intuitive type-hierarchy for events is why other criteria are needed.
   (b) The presuppositions, expectations, and concomitants of the targets may differ. For example, *cross-examine* evokes a courtroom session, a preceding event of questioning by an opposing legal party, etc., differentiating it from the simpler *examine*. By this feature, *receive* and *take* would be differentiated, since *receive* presupposes another willing agent participating as the Donor and *take* does not.
2. For semantic roles:
   (a) Their number and type, (e.g. *kill* has a role not present for *die*)
   (b) Interrelations (e.g. Purpose refers to a particular kind of mental state of an Agent, as opposed to Means which refers to an action of an Agent)
   (c) Profiling (e.g., the difference of *buy* and *sell*, in which both refer to a Buyer and a Seller, but in the case of *buy* the Buyer is portrayed as more saliently controlling the action, vs. *sell*, in which the Seller is portrayed as more salient), and
   (d) The semantic preferences and restrictions the target imposes (e.g. *tie* requiring the Connector be a long, flexible object).

[1] This neglects two other senses, unseen in the Wall Street Journal: 'remove the stems from something which inherently has a stem' as in *John stemmed the cherries* and 'reduce something down to just a stem' as in a morphological stemmer/lemmatizer.

Grouping usages according to close matches of such features allows FrameNet to form "minimal" frames; the more-inclusive final frames are then formed by loosening some of these conditions such as 2d., so that *tie* and *staple* can be grouped despite the constraints on what kind of Connector they specify.

Conversely, these semantic considerations (especially 2b.) led FrameNet to draw a distinction between causative and inchoative cases that PropBank does not make. Lexical membership in a FrameNet frame entails that for each use of a target, all of the core frame elements must be semantically present. Inchoatives do not entail the existence of a Cause or Agent, as can be seen by comparing *the rain ended* to the infelicitous ?*(someone/something) ended the rain.* The inchoative and causative uses of *end* thus belong to the frames Process_end and Cause_to_end respectively.

4. Consequences in the Annotation

The different aims of PropBank, FrameNet and SALSA are reflected in the practice of annotation. PropBank limits itself to annotating the literal meaning of the target, preferring small, incremental, easily-attained goals. FrameNet and SALSA follow Fillmore (1985), which states that *'Frame Semantics does not seek to draw an a priori distinction between semantics proper and (an idealized notion of) text understanding'* and consequently try to annotate what is actually understood. This makes the task more complex but should finally yield a more informative annotation.

Semantic annotation has to deal with large classes of phenomena for which the meaning is either hard to pin down or subject to debate. We now show the consequences of different annotation choices of the three frameworks for such phenomena for both tasks of frame(set) assignment and semantic role assignment. For the first task, we discuss metaphors, support cases, and instances with multiple meaning aspects, while issues for the second task are phrases realizing multiple semantic roles and nonlocal semantic roles.

4.1. Metaphor

Metaphors are abundant even in newspaper texts. A recent study of a 100k word corpus found that roughly 54% of all motion terms were used metaphorically (Tewari, 2003). (5) is a case in point.

(5) Viele meinen, dass Perot mit seinem Befehlston auf dem Capitol gegen eine Wand **laufen** würde. (Tiger s13)
(Many think that Perot would **walk** into a brick wall on the Capitol with his commanding tone.)

In such cases, annotation projects have to decide between annotation the 'source' (literal) or 'target' (metaphorical) meaning (following Lakoff and Johnson's (1980) terminology). However, the border between metaphor proper, and cases that are lexicalized so far as to be indiscernible as a metaphor, is often not clear-cut, as in (6). *Get through [a difficult time]* could be characterized as a metaphor with a Motion source, but can also be seen as lexicalized so far to have become a separate sense of *get*.

(6) Der "Pluralismus von Erklärungen" aus der CDU/CSU-FDP-Koalition zeige, dass die Einigkeit über die Pflegeversicherung nur "vorgetäuscht" worden sei, "um über die Sommerpause zu **kommen**", sagte Klose.
(The "multiplicity of explanations" given by the CDU/CSU coalition showed that they only "pretended" to agree on nursing care insurance "in order to **get through** the summer break", Klose said.)

PropBank. PropBank, for the most part, takes a consistently literal analysis of such constructions. A later pass of annotation is planned, in which instances will be flagged as being metaphorical. Nevertheless, there are cases when metaphor is unavoidable. These tend to occur with the most frequent verbs, those with the most bleached underlying semantics. When these are common enough, they can be thought of as being just another sense of the verb and thus acquire a new frameset. The division between a true metaphor and a different sense is not clear, however: how often is often enough?

FrameNet. FrameNet decides between conventionalized metaphors, like (5), and nonce metaphors, such as in (7), whose unique meaning is determined by its special context. Conventionalized metaphors are annotated with the target frame, while nonce metaphors are ignored, or in rare cases they are annotated and tagged with the sentence-level tag "Metaphor".

(7) A small gust of laughter **blew** through him , and left him smiling . (BNC)

SALSA. In the finished SALSA corpus, both the source and the target frame will be assigned. To speed up annotation, however, the tagging of metaphoric instances is split up into two passes. In the current first pass, the instance in question is marked as metaphoric, and either the source or the target frame is tagged (with a flag indicating which of the two it is). The annotators mark whichever of the two frames is easier to determine; the default is the source, since the target meaning is sometimes hard to pin down in terms of frames. (8) shows such a case.

(8) Den einen **geht** der Schritt **zu weit**, den anderen nicht **weit genug**. (TIGER s10471)
(For some this **goes too far**, for some, not **far enough**.)

(8) talks about some cognitive scale, maybe one of acceptability. But the target sense can only be described on a very abstract level, much more abstract than is usual in frame descriptions.

4.2. Support constructions

Support constructions are non-compositional multiword expressions[2] in which a governing verb and/or preposition are optional for lending the phrase, semantically headed by

[2]Non-compositionality is tested by substitutability, replacing the words of the phrase with likely synonyms. If the synonyms do not allow the phrase to retain a similar meaning, then it is non-compositional and should be annotated as a unit.

a noun or adjective, its essential meaning. Putting it slightly more formally, a support construction involves (1) an adjective or noun that denotes a state or event and is the source of virtually all the meaning of the phrase and (2) syntactically governing verbs or prepositions with only simple, grammatical meaning which do not have the same meaning independently of the target.

The simplest cases are phrases like *take a bath*, which evokes the Grooming frame, in which *bath* (as in *his bath lasted three hours*) all by itself evokes the exact same frame. *Be in possession (of)* provides another clear, but slightly more complicated example. Here, *be* and *in* are supports, because when we compare *John is in possession of the estate* and *John's possession of the estate*, the differences in meaning are not framal differences.

One obvious problem that the idea of support presents for any semantic annotation project is how to recognize and record the cases, and how to record the differences between cases. A further basic problem is what types of 'minor' meaning change are allowed for the supports themselves, such as causativity, aspectual change, etc, and how to record the differences between them.

For support, as for the other phenomena we have discussed, there are borderline cases that could be characterized as support as well as something else. This problem occurs particularly often with high-frequency verbs that can denote situatedness, like *put*, *lie*, *stand*. The trouble with cases like (9) is that they could be analyzed either as a simple case of support, or as a metaphor with the frame Being_situated as a source.

(9) Zwar **liege** die Verantwortung allein bei der Bundesregierung , doch angesichts der nicht unerheblichen Gefährdung der eingesetzten Soldaten habe man eine breite Zustimmung gesucht, sagte ein Sprecher. (TIGER s1307)
(While responsibility **lies** solely with the federal government, broad agreement had been sought in view of the considerable danger for the soldiers, a spokesman said.)

FrameNet. The types of meaning change allowed by FrameNet for supports include:

Vanilla: the support adds virtually nothing to the target (like the *take a bath* example above).

Aspectual: the support changes the temporal focus of the event portrayed by the target, e.g. *get/go/fall into a (foul) mood* vs. (the vanilla) *be in a (foul) mood*.

Point-of-view: the support changes the profiled point-of-view of the target, e.g. *undergo* in *undergo a physical exam* vs. *give a physical exam*, with the patient's and doctor's points of view respectively.

Causative: the support adds another participant and the idea of causation of the basic scene. These generally occur paired with a non-causative support, as in *put in a (foul) mood* vs. *be in a (foul) mood*; *bring into play* vs. *come into play*; *give a headache* vs. *have a headache*, or the idiosyncratic *show a good time* vs. *have a good time*.

Currently, FrameNet annotates supports with a special tag, and only when they occur in the context of a noun or adjective target that is already being examined. There is currently no annotation of supports as targets themselves, and no systematic way of recognizing instances of the separate types given above.

PropBank. PropBank dodges the entire issue by lumping all support constructions for each verb into a single frameset, described as 'support'. These framesets usually take two or three roles, of which one is the noun which is the real predicate and the others are the roles of the nominal. For example,

(10) [$_{\text{Arg0}}$ John] **made** [$_{\text{Arg1}}$ a shrewd guess about Mary's intentions].

For those cases where the predicate nominal is deverbal, the Nombank project at New York University is annotating the semantic role structure using the PropBank lexical frames, so a sentence such as (10) will receive a second, overlapping structure:

(11) [$_{\text{Arg0}}$ John] made a [$_{\text{ArgM-MNR}}$ shrewd] **guess** [$_{\text{Arg1}}$ about Mary's intentions].

SALSA. In the first pass, SALSA tags all the cases recognized by FrameNet above just as the Pseudo-frame Support. This is somewhat similar to the PropBank treatment. A second pass over the support cases is planned, giving them a deeper, FrameNet-style analysis.

4.3. Words with several simultaneous meaning components

There are words with several simultaneous meaning components, which are unlike polysemy in that the different meanings are not a question of context, but rather refer to two simultaneous situations at once. This can be either restricted to a single instance or systematic. For example, many verbs can be systematically used to describe both the content of a communication and its manner:

(12) And don't expect many complete games by pitchers – perhaps three out of 288, **laughs** *t* Mr. Fingers, the former Oakland reliever. (wsj_0214)

The following idiosyncratic case demonstrates that such cases can show a blend of the syntactic patterns of the two single usages. (13) has both the direct speech of a "communication", and the direct object of the "impede" meaning:

(13) "Sorry, you cannot enter", he **blocked** the way.

The question of how to annotate these cases has an obvious impact on the usefulness of the annotation, since in order to be aware of the full meaning potential of the expression, one would need to indicate both (or all) components.

PropBank. To annotate an instance such as (12), PropBank creates a new frameset for *laugh*; while the main frameset includes only a single role for the laugher, this new frameset must also include a role for the utterance spoken while laughing. Since this behavior is systematic for 'manner of speaking' verbs (including *laugh, cry, wheeze*

and many others), this policy can lead to a proliferation of framesets. The same is true for idiosyncratic cases such as (13).

FrameNet. In FrameNet, blended frames are constructed for targets that systematically exhibit several meaning aspects, like (12), while idiosyncratic cases such as (13) are not treated[3].

SALSA. In SALSA, instances with multiple meaning aspects can be marked with more than one frame, in accordance with the general underspecification principle used in SALSA annotation. For (12), the applicable frames would be Statement and something like Physical_obstruction.

A particularly difficult case arises when a sentence might be seen to evoke multiple senses or not, depending on the view of the reader. Unlike the cases above, where the multiple senses are clearly present, the senses available in the sentences below are much more subtle and optional:

(14) Such a thought would never **cross my mind**.

(15) I must admit I feel a tad embarrassed though, as the idea of focusing on the local market first didn't even **cross my mind**... (www.webhostingtalk.com/archive/thread/232858-1.html, Feb 24, 2004)

Literally this construction means that the speaker would not think of X, that X would never occur to him or her. But (14) also has the connotation of not wanting to do (the thing which was referred to). (15), seemingly identical to (14), seems to not invoke this secondary sense, apparently related to the lack of the modal *would*. Both examples, (14) and (15), share the idea of Invention, and the first example also includes the idea of Desiring. How many of these senses should annotators mark for each of these sentences? Should annotators simply tag (14) as a case of Invention, or should the secondary sense also be indicated?

The annotation of these examples is unproblematic in PropBank since there are no syntactic peculiarities. FrameNet would differentiate the two examples with one target *would cross mind* and another simply *cross mind*. The first would be in a frame which inherits from Desiring, the second would not. SALSA treats such borderline examples on a case-by-base basis, letting annotators decide between single-frame annotation and underspecification on the basis of the prevailing overall meaning of the sentence.

4.4. Phrases realizing multiple semantic roles

We now turn to phenomena that concern semantic role assignment rather than frame(set) assignment. The first phenomenon parallels target words with multiple meaning aspects: lexical material that simultaneously bears multiple semantic roles. This situation often arises with a plural constituent within which two separate semantic roles have been merged, as the contrast between (16) and (17) shows:

(16) Argentine negotiator Carlos Carballo will **meet** with banks this week. (wsj_0021)

(17) The economic and foreign ministers of 12 Asian and Pacific nations will **meet** in Australia next week to discuss global trade as well as regional matters such as transportation and telecommunications. (wsj_0043)

Sentence (18) is much more complex. The expression *under the hand dryer* is certainly the Place role of the drying, but it also indicates the Instrument of the drying, the hand dryer.

(18) We immediately rushed to the ladies, washed Jessica carefully in the sink and **dried** her under the hand dryer. (BNC)

Note that in this example a certain amount of knowledge about hand dryers (namely that they usually blows hot air downward) is required of the annotator, as well as some degree of inference, in determining that the Place doubles as the Instrument of the drying. Note also that the assignment of Instrument to *under the hand dryer* is defeasible and can be overwritten, for example by continuing the sentence by *... using lots of paper towels*.

PropBank. Under PropBank there is no provision for a single constituent to bear multiple labels, so the annotators are forced to choose. For these and similar cases a hierarchical notion of semantic roles was developed, preferring lower-numbered to higher-numbered labels and numbered labels to ArgMs, which are felt to be universal and adjunct-like. In (17), with a choice between Arg0 and Arg1, the lower numbered label, Arg0, is used, and in (18), the instrumental Arg2 is used in preference to the location ArgM-LOC. While annotators have experienced little difficulty with this policy, it might pose interesting challenges to systems attempting to interpret PropBank annotations.

FrameNet/SALSA. In FrameNet, *meet* in the sense of (17) is in the Discussion frame, which has one role for the subject participant when it refers to the collective Interlocutors, and two other roles (Interlocutor_1 and Interlocutor_2) for the subject and complement respectively when these denote participants separately. The relationships between the roles (Interlocutor_1 requires Interlocutor_2 and Interlocutors excludes Interlocutor_1 and Interlocutor_2) are specifically encoded in the database. Sentence (18) is annotated for both semantic roles, Place and Instrument, in the FrameNet corpus.

4.5. Nonlocal semantic roles

Since relatively few sentences, especially in more formal or journalistic registers, contain only one clause, the question of the scope of annotation often arises. How far away from a verbal head does one look for roles of that verb? Interestingly, roles that are realized nonlocally show the same characteristics as the Instrument role in (18): World knowledge, as well as some inference, is required to a much larger degree than usual to assign these semantic roles, and the assignment is defeasible, i.e. it is possible to change the way semantic roles are assigned by setting the expressions in

[3]One possibility of analyzing (13) would be to annotate block.v just for Physical_obstruction, and to introduce an extrathematic Message role to that frame's definition. Like all extrathematic roles, it would be introduced by some kind of construction with its own separate semantics and form constraints – i.e. it is not introduced by the target, but the target can unify with it.

question in a different context. For example:

(19) Besitzer von Zweifamilienhäusern, die vor 1987 gebaut oder **gekauft** haben... (TIGER s975)
(Owners of two-family homes who have built or **purchased** before 1987)...

Gekauft (purchased) evokes the frame Commerce_sell. *Zweifamilienhäusern* (two-family homes), which is not in the maximal projection of *gekauft*, may be inferred to be the Goods role of this frame. This inference is defeasible, though. Suppose this sentence occurs within a text about buying stocks. Then the Goods may be the stocks instead of the houses.

Noun targets are especially problematic in that most of their roles are usually realized nonlocally and are therefore defeasible, such as in (20):

(20) Vor Jahren, als Helmut Kohl erstmals ganz unten war [...], machte [...] Günter Oettinger bundesweit mit einer Rücktritts**forderung** von sich reden. (TIGER s1862)
(Years ago, when Helmut Kohl was on the rocks for the first time, Günter Oettinger brought himself into public awareness with a **demand** for resignation.)

Forderung (demand) evokes the frame Request. Neither the Speaker nor the Addressee of the request are realized locally. The Speaker is probably Günter Oettinger, but he might also be just a medium. The Addressee is probably Helmut Kohl; nevertheless, all these inferences can be overridden by context.

FrameNet FrameNet allows annotators to annotate nonlocal arguments when they participate in any of a number of recognized nonlocal constructions such as questions and fronting, or in general when we can recognize that the assignment of the semantic roles is not defeasible by context.

SALSA. In SALSA nonlocal semantic roles are included in the annotation, for three reasons: First. annotators usually have strong intuitions about these nonlocal semantic roles and tend to annotate them when this is not explicitly prohibited. Second, these nonlocal, defeasibly inferred semantic roles constitute interesting data on inferences people make when understanding sentences. Third, local and nonlocal semantic roles can be clearly distinguished through the syntactic structure, which makes it possible to sort out nonlocal roles whenever that is required.

PropBank. PropBank, being built upon the existing syntactic parse in the Penn English Treebank II, makes use of the 'traces' (overt markers on empty nodes, coindexed with their lexical antecedents) present in the treebank to find nonlocal arguments.

There are cases, however, when there is a genuine ambiguity as to the antecedent of a trace, such as in (21):

(21) Commonwealth Edison is seeking about $245 million in rate increases [*T*-1] to **pay** for Braidwood 2. (wsj_0015)

In this example, the trace [*T*-1] could point to *Commonwealth Edison*, who will be doing the paying after all, or to *$245 million in rate increases* which is the instrument of paying. There are many cases of this agent/instrument ambiguity in trace chains, leading PropBank to choose the agent in all cases.

5. Conclusion

Whether in handling metaphors, identifying support cases, or assigning a single sense to a role-bearing expression, all three projects have to deal with corpus instances that lie on the borderline between different categories and defy clear classification. Our comparison has shown that the theoretical differences between the three projects lessen in view of actual annotation, however the mechanisms the three projects use in dealing with borderline cases differ. PropBank tends to formulate general policies (e.g. preferring arguments with lower numbers when more than one role label applies to a phrase), FrameNet includes systematic cases, but excludes idiosyncratic borderline cases from its consideration, and SALSA allows for more than one tag through underspecification, allowing for a later analysis based on annotated underspecified instances.

One deeper question that our comparison has highlighted is: How much context information should annotators use in determining the tag to be assigned, and how much inference are they allowed to perform to divine the meaning of an expression? This question is probably most prominent in the *cross one's mind* example (14) for frame(set) assignment, and for semantic role assignment the *hand dryer* example (18) and the nonlocal cases. While all three projects have to allow some use of context in determining the meaning of a phrase, FrameNet and SALSA have to take it into account to a larger degree since they are assigning the meaning that is actually understood, while PropBank mostly focuses on the literal meaning.

6. References

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

Burnard, L., 1995. *User's guide for the British National Corpus*. British National Corpus Consortium, Oxford University Computing Services.

Erk, K., A. Kowalski, S. Pado, and M. Pinkal, 2003. Towards a resource for lexical semantics: A large German corpus with extensive semantic annotation. In *Proceedings of ACL-03*. Sapporo, Japan.

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

Johnson, C. R., C. J. Fillmore, M. R. L. Petruck, C. F. Baker, M. J. Ellsworth, J. Ruppenhofer, and E. J. Wood, 2002. FrameNet: Theory and Practice. `http://www.icsi.berkeley.edu/~framenet/book/book.html`.

Kingsbury, P. and K. Kipper, 2003. Deriving verb-meaning clusters from syntactic structure. In *Workshop on Text Meaning, HLT/NAACL 2003*. Edmonton, Canada.

Kingsbury, P. and M. Palmer, 2002. From TreeBank to PropBank. In *Proceedings of LREC-2002*. Las Palmas, Spain.

Kipper, K., M. Palmer, and O. Rambow, 2002. Extending PropBank with VerbNet semantic predicates. In *Workshop on Applied Interlinguas, held in conjunction with AMTA-2002*. Tiburon, CA.

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

Levin, B., 1993. *English Verb Classes and Alternations*. Chicago, IL: University of Chicago Press.

Marcus, M., G. Kim, M.A. Marcinkiewicz, R. MacIntyre, A. Bies, M. Ferguson, K. Katz, and B. Schasberger, 1994. The Penn Treebank: Annotating predicate argument structure. In *ARPA Human Language Technology Workshop*. Plainsboro, NJ.

Tewari, A., 2003. Detecting metaphors involving motion verbs in online business news sources. `www.eecs.berkeley.edu/~ambuj/research/reports/cs182report.pdf`.