'Deep' Grammatical Relations for Semantic Interpretation

Mark McConville and **Myroslava O. Dzikovska**
Institute for Communicating and Collaborative Systems
School of Informatics, University of Edinburgh
Informatics Forum, 10 Crichton Street, Edinburgh, EH8 9AB, Scotland
{Mark.McConville,M.Dzikovska}@ed.ac.uk

Abstract

In this paper, we evaluate five distinct systems of labelled grammatical dependency against the kind of input we require for semantic interpretation, in particular for the deep semantic interpreter underlying a tutorial dialogue system. We focus on the following linguistic phenomena: passive, control and raising, noun modifiers, and meaningful vs. non-meaningful prepositions. We conclude that no one system provides all the features that we require, although each such feature is contained within at least one of the competing systems.

1 Introduction

The aim of the work reported in this paper is to evaluate the extent to which proposed systems of grammatical relations (GRs) reflect the kinds of deep linguistic knowledge required for semantic interpretation, in particular for deriving semantic representations suitable for domain reasoning in dialogue systems.

Grammatical relations either produced by or extracted from the output of wide-coverage syntactic parsers are currently used as input to shallow semantic parsers, which identify semantic relations that exist between predicators (typically verbs) and their dependents (Gildea and Jurafsky, 2002; Erk and Padó, 2006). Predicate-argument structure identified in this way can then be used in tasks like information extraction (Surdeanu et al., 2003) and question answering (Kaisser and Webber, 2007).

However, wide-coverage stochastic parsers are only rarely used in dialogue systems. Traditionally, interpretation modules of dialogue systems utilise specialised parsers and semantic interpreters handcrafted to a small domain (Seneff, 1992; Chang et al., 2002), or wide coverage deep parsers (Allen et al., 2007; Jordan et al., 2006; Wolska and Kruijff-Korbayová, 2003; Callaway et al., 2007; Kay et al., 1994). Unlike in information retrieval and question answering tasks, the system often needs to be connected to a knowledge base which represents the state of the world, and must be able to convert user utterances into knowledge base queries. In addition to identifying predicate-argument relationships, such systems need to support a variety of tasks, for example resolution of pronouns and anaphors, and interpreting negation, quantification, tense and modality.

While deep parsers produce precise semantic representations appropriate for such reasoning, they suffer from robustness problems. Wide-coverage dependency parsers could potentially provide a more robust alternative, provided that their output is easy to convert into semantic representations for reasoning.

Section 2 introduces the kind of deep linguistic processing application which motivates our approach to grammatical relations. Section 3 defines some underlying principles behind the kind of 'deep' GR system we have in mind. The remainder of the paper discusses a number of linguistic phenomena in detail, and evaluates how well various systems of GR representation from the dependency parsing literature capture the kind of linguistic insights required for interface with reasoning — passive (section 4), raising and control (section 5), noun modification (section 6) and syntactic versus semantic prepositions (section 7).

2 Motivation

As an example application that requires deep parsing consider a tutorial dialogue system that interprets students' answers to factual questions (e.g. *Which bulbs will be lit in this circuit?*) as well as explanation questions (e.g. *Explain your reasoning!*). It has been argued previously (Wolska and Kruijff-Korbayová, 2004; Rosé et al., 2003) that tutorial dialogue systems require deep understanding of student explanations, which can have significantly more complex structure than database queries in the information-seeking domain. In our application, if a student is asked for an explanation, his or her input has to be passed through the domain knowledge base to verify its factual correctness, and a separate process verifies that all relations mentioned in the explanation are correct and relevant. For example, imagine that the student says the following:

(1) The bulbs in circuits 1 and 3 will be lit because they are in closed paths with the batteries.

Here, the system has to verify two things: (a) that the facts are correct (bulbs in circuits 1 and 3 will be lit, and each of those bulbs is in a closed path with a battery); and (b) that the reason is valid — being in a closed path with a battery is a necessary and sufficient condition for a bulb to be lit.

This task is particularly interesting because it combines characteristics of deep and shallow interpretation tasks. On the one hand, the fact-checking mechanism requires a connection to the database. Thus, both pronouns and definite noun phrases need to be resolved to the objects they represent in the knowledge base, and first-order logic formulas representing utterance content need to be checked against the system knowledge. This task is similar to natural language interfaces to databases, or knowledge acquisition interfaces that convert language into knowledge base statements (Yeh et al., 2005). On the other hand, with respect to reason checking, human tutors have indicated that they would accept an answer simply if a student produces the key concepts and relations between them, even if the answer is not strictly logically equivalent to the ideal answer (Dzikovska et al., 2008). Human tutors tend to be especially lenient if a student is asked a generic question, like *What is the definition of voltage?*, which does not refer to specific objects in the knowledge base. Thus, a simpler matching mechanism is used to check the reasons, making this task more similar to an information retrieval task requiring shallower processing, i.e. that the predicate-argument relations are retrieved correctly (though negation still remains important).

Thus, while a specific task is used to motivate our evaluation, the conclusions would be applicable to a variety of systems, including both deep and shallow semantic interpreters.

For the purposes of this evaluation, we discuss features of grammatical representation relevant to two subtasks critical for the system: (a) identifying predicate-argument structure; and (b) resolving anaphora.

The extraction of predicate-argument relations is a common requirement for both shallow and deep semantic tasks. For example, for the student input in example (1) we may expect something like:[1]

(2)
```
(LightBulb b1) (LightBulb b2)
(lit b1 true) (lit b2 true)
(Path P3) (closed P3 true)
(contains P3 b1) (Path P4)
(closed P4 true) (contains P4 b2)
```

Resolving anaphora, on the other hand, is particularly important for the kind of deep semantic processing used in dialogue systems. Implicit in the above representation is the fact that the definite noun phrase *the bulbs in circuits 1 and 3* was resolved to domain constants `b1` and `b3`, and indefinite references to paths were replaced by Skolem constants `P3` and `P4`. The reference resolution process requires detailed knowledge of noun phrase structure, including information about restrictive modification, and this is the second focus of our evaluation.

Ideally, we would like a dependency parser to produce grammatical relations that can be converted into such semantic representations with minimal effort, thus minimising the number of specific rules used to convert individual relations. We discuss the principles underlying such representations in more detail in the next section.

---

[1] We used a simplified representation of quantifiers that assumes no scope ambiguity and uses skolem constants to represent existential quantification. This is sufficient for our particular application. In general, a more sophisticated quantifier representation would be necessary, for example that proposed in Copestake et al. (2005) or Bos and Oka (2002), but we leave the relevant evaluation for future work.

3 Deep grammatical relations

We formulated **four** principles for deep grammatical relations representation.

Firstly, grammatical relations should, whenever possible, reflect relations between the predicators (i.e. content words as opposed to function words) in a sentence. In addition, the same relation should correspond to the same role assignment. For example, the deep GRs in passive constructions should be the same as those in the active equivalents (see section 4), and the analysis of a control verb construction like *John persuaded Mary to dance* should make it clear that there is a 'subject' GR from *dance* to *Mary* similar to that in the implied sentence *Mary danced* (see section 5).

Secondly, a GR should, whenever possible, appear only if there is a an explicit selectional restriction link between the words. For example, in a raising verb construction like *John expects Mary to dance*, there should be **no** GR from the raising verb *expects* to its object *Mary* (see section 5). Also, where a preposition functions strictly as a syntactic role marker, as in the construction *John relies on Mary*, it should have no place in the GR analysis; rather there should be a direct link from the verb to the embedded noun phrase (see section 7).

Thirdly, the GRs should preserve evidence of syntactic modification to enable reference resolution. To understand why this is important, take the following two examples:

(3)   The lit bulb is in a closed path.
      The bulb in a closed path is lit.

From a pure predicate-argument structure perspective, these two sentences share exactly the same deep GRs:[2]

(4)   `ext(lit,bulb)`
      `ext(in-closed-path,bulb)`

However, from the perspective of reference resolution, the two sentences are very different. For the first example, this process involves first finding the lit bulb and then verifying that it is in a closed path, whereas for the second we need to find the bulb in a closed path and verify that it is lit. This difference can be captured by assigning the following additional deep GRs to the first example:

[2]The representation is simplified for reasons of exposition. The GRs should be interpreted as follows: `ext` denotes the external argument of an adjective or preposition, `ncmod` a non-clausal restrictive modifier, and `det` the determiner of a noun.

(5)   `det(bulb,the)`
      `ncmod(bulb,lit)`

And the following GRs are added to the analysis of the second example:

(6)   `det(bulb,the)`
      `ncmod(bulb,in-closed-path)`

Now the two analyses are formally distinct: (a) the first is rooted at predicate *in a closed path* and the second at *lit*; and (b) the definite external argument *the bulb* takes scope over the modifier *lit* in the first but over *in a closed path* in the second. Noun modification is discussed in section 6.

Finally, the set of grammatical relations should make it easy to identify and separate out constructions which are largely dependent on semantic/world knowledge, such as N-N modification, so that separate models and evaluations can be conducted as necessary.

4 Passive

The shared task dataset contains numerous passive participles, most of which can be classified into the following four groups depending on how the participle is used: (a) complement of passive auxiliary e.g. *Tax induction is **activated** by the RelA subunit*; (b) complement of raising verb e.g. *The administration doesn't seem **moved** by the arguments*; (c) nominal postmodifier e.g. *the genes involved in T-cell growth*; and (d) nominal premodifier e.g. *the proposed rules*.

In all these cases, our system for deep grammatical relation annotation requires: (a) that there is a relation **from** the passive participle **to** the deep object; and (b) that this relation be the same as in the corresponding active declarative construction, so that predicate-argument structure can be straightforwardly derived. Thus, for example, the analysis of *Tax induction is **activated** by the RelA subunit* will contain the GR `dobj(activated,induction)`, and that of *the proposed rules* will include `dobj(proposed,rules)`, where `dobj` is the relation between a transitive verb and its (deep) direct object.

We evaluated five GR-based output formats according to these two features. The results are presented in Table 1, where for each representation format (the rows) and each usage class of passive participles (the columns), we provide the GR which goes **from** the participle **to** its deep object,

	complement of passive auxiliary	complement of raising verb	nominal postmodifier	nominal premodifier	active
HPSG	ARG2 (of verb_arg12)				
RASP	ncsubj:obj				dobj
CCGBank	Spss\NP			N/N	S\NP/[NP]
Stanford	nsubjpass		-		dobj
PARC	subj		-		obj

Table 1: Representation of deep objects in passive and active

if such a GR exists.[3] The five GR representations compared are:

HPSG predicate-argument structures extracted from the University of Tokyo HPSG Treebank (Miyao, 2006)

RASP grammatical relations as output by the RASP parser (Briscoe et al., 2006)

CCGBank predicate-argument dependencies extracted from CCGBank (Hockenmaier and Steedman, 2007)

Stanford grammatical relations output by the Stanford Parser (de Marneffe et al., 2006)

PARC dependency structures used in the annotation of DepBank (King et al., 2003)

The first four columns in Table 1 represent, for each of the four uses of passive participles listed above, the grammatical relation, if any, which typically joins a passive participle to its deep object. The rightmost column presents the label used for this relation in equivalent active clauses. Adjacent columns have been collapsed where the same GR is used for both uses. The ideal system would have the same GR listed in each of the five columns.

The grammatical relations used in the Stanford, PARC and RASP systems are atomic labels like `subj`, `obj` etc, although the latter system does allow for a limited range of composite GRs like `ncsubj:obj` (a non-clausal surface subject which realises a deep object). In the HPSG system, verbal subjects and objects are represented as `ARG1` and `ARG2` respectively of strict transitive verb type `verb_arg12`. Finally, the GRs assumed in CCGBank consist of a lexical category (e.g. the strict transitive verb category `S\NP/NP`) with one argument emphasised. I assume the

following notational convenience for those categories which contain specify more than one argument — the emphasised argument is surrounded by square brackets. Thus, subject and object of a strict transitive verb are denoted `S\[NP]/NP` and `S\NP/[NP]` respectively.

With respect to Table 1, note that: (a) in the CCGbank dependency representation, although prenominal passive participles *are* linked to their deep object (i.e. the modified noun), this relation is just one of generic noun premodification (i.e. `N/N`) and is thus irrelevant to the kind of predicate-argument relation we are interested in; (b) in the PARC and Stanford dependency representations, there is no GR from noun-modifying passive participles to their deep objects, just generic modification relations in the opposite direction; and (c) in PARC, passive participles are themselves marked as being passive, thus allowing a subsequent interpretation module to normalise the deep grammatical relations if desired.

If we are interested in a system of deep grammatical role annotation which allows for the representation of normalised GRs for passive participles in all their uses, then the HPSG Treebank format is more appropriate than the other schemes, since it uniformly uses deep GRs for both active and passive verb constructions. The RASP representation comes a close second, only requiring a small amount of postprocessing to convert `ncsubj:obj` relations into `dobj` ones. In addition, both the CCGBank and the Stanford notation distinguish two kinds of surface subject — those which realise deep subjects, and those which realise passivised deep objects.

5 Control

The shared task dataset contains a number of infinitives or participles which are dependents of non-auxiliary verbs or adjectives (rather than being noun modifiers for example). Most of these can

---

[3]The relations presented for HPSG and CCG are those for passive participle of *strict* transitive verbs.

	complements	adjuncts	raising
HPSG	✓	✓	✗
RASP	✓	✓	✗
CCGbank	✓	✓	✗
Stanford	✓	✗	✓
PARC	✗	✗	✗

Table 2: Representation of controlled subjects and raising

be partitioned into the following three classes: (a) complements of subject control verbs e.g. *The accumulation of nuclear c-Rel **acts** to **inhibit** its own continued production*; (b) complements of subject raising verbs e.g. *The administration **seems moved** by arguments that …*; and (c) subject controlled adjuncts e.g. *Alex de Castro has stopped by to **slip** six cards to the Great Man Himself.*

In all these cases, our deep grammatical role annotation requires that there be a subject relation (or an object relation in the case of a passive participle) from the infinitive/participle to the surface subject (or surface object in the case of object control) of the controlling verb/adjective. For example, the analysis of *Tax acts indirectly by inducing the action of various host transcription factors* will contain both the GRs `sbj(acts,Tax)` and `sbj(inducing,Tax)`. In addition, we also want to distinguish 'raising' verbs and adjectives from control structures. Thus, in the analysis of *The administration **seems moved** by arguments that …*, we want a (deep) object relation from *moved* to *administration*, but we don't want **any** relation from *seems* to *administration*.

We again evaluated the various GR-based output formats according to these features. The results are presented in Table 2, where for each representation format (the rows) we determine: (a) whether a verb with an understood subject which is a *complement* of the matrix verb is linked directly to its relevant subject (column 1); (b) whether a verb with an understood subject which is a controlled *adjunct* of the matrix verb is linked directly to its relevant subject (column 2); and (c) whether raising verbs are non-linked to their surface subjects (column 3). Note that the Stanford dependency representation is the only format which distinguishes between raising and control. This distinction is made

both structurally and in terms of the name assigned to the relevant dependent — controlled subjects are distinguished from all other subjects (including raised ones) by having the label `xsubj` rather than just `nsubj`.[4]

The ideal GR representation format would have a tick in each of the three columns in Table 2. It is clear that no single representation covers all of our desiderata for a deep grammatical relation treatment of control/raising, but each feature we require is provided by at least one format.

6 Nominal modifiers

The dataset contains numerous prenominal modifiers[5], subdivided into the following three groups: (a) attributive adjectives e.g. *a few **notable** exceptions*; (b) verb participles e.g. *the **proposed** rules*; and (c) nouns e.g. *a **car** salesman.*

In order to ensure an adequate representation of basic predicate-argument structure, our system of deep grammatical annotation first of all requires that, from each prenominal adjective or verb, there is an appropriate relation **to** the modified noun, of the same type as in the corresponding predicative usage. For example, assuming that *He proposed the rules* has a direct object relation from *proposed* to *rules*, the same relation should occur in the analysis of *the proposed rules.* Similarly, if *The exceptions are notable* is analysed as having an external argument relation from *notable* to *exceptions*, then the same should happen in the case of *a few notable exceptions.* However, this does not appear to hold for prenominal nouns, since the relation between the two is not simply one of predication — *a car salesman* is not a salesman who 'is' a car, but rather a salesman who is 'associated' with cars in some way. Thus we would not want the same relation to be used here.[6]

Secondly, in order to ensure a straightforward interface with reference resolution, we need a modification relation going in the opposite direc-

---

[4]We have judged that CCGBank does not make the relevant distinction between raising and control verbs based on the dependency representations contained in the shared task dataset. For example, for the example sentence *The administration seem moved by the fact that …*, a CCG subject relation **is** specified from the raising verb *seem* to its surface subject *administration.*

[5]We focus on prenominal modifiers in order to keep the exposition simple. Similar remarks are valid for postnominal restrictive modifiers as well.

[6]Presumably the same goes for attributive adjectives which lack corresponding predicative uses, e.g. *the **former** president.*

tion, **from** the modified noun **to** each (restrictive) modifier, as argued in section 2. Thus, a complete GR representation of a noun phrase like *notable exceptions* would be cyclical, for example:

```
(7)  ext(notable,exceptions)
     ncmod(exceptions,notable)
```

We evaluated the various GR-based output formats according to these desiderata. The results are presented in Table 3. For each annotation scheme (the rows), we first present the relation (if any) which goes **from** the modified noun **to** each kind of pre-modifier (adjective, verb participle and noun respectively).[7] The middle three columns contain the relation (if any) which goes **to** the noun **from** each kind of modifier. Finally, the last three columns give the corresponding predicative relation used in the annotation scheme, for example in constructions like *The exceptions are notable*, *He proposed the rules*, or *Herbie is a car*. Where it is unclear whether a particular format encodes the relation between a predicative noun and its subject, we mark this as '?' in the last column.

Ideally, what we want is a representation where: (a) there is a GR in all nine columns (with the possible exception of the 'noun modifier to noun' one (column 6)); (b) the corresponding relations in the middle and righthand sections are identical, *except* for 'noun modifier to noun' (column 6) and 'predicative noun' (the last column) which should be distinct, since the relation between a noun modifier and its head noun is not simply one of predication.

It is clear that no one representation is perfect, though every feature we require is present in at least one representation system. Note in particular that the HPSG, PARC and Stanford systems are acyclic — the former only has 'modifier to noun' links, while the latter two only have 'noun to modifier' ones. The RASP format is cyclic, at least for prenominal participles — in *the proposed rules*, there is a modifier relation from *rules* to *proposed*, as well as a deep object relation from *proposed* to *rules*, the same relation that would be found in the corresponding predicative *the rules were proposed*.

Note finally that the PARC and Stanford representations distinguish between prenominal adjectives and nouns, in terms of the name of the relevant modifier GR. This corresponds well with our

---

[7]Note that the N/N links in the CCG representation actually go from the modifier to the noun. However, they have been included in the set of 'noun to modifier' relations since they are formally modifier categories (i.e. of the form X/X).

preference for a GR system where we can evaluate modules of N-N disambiguation (e.g. *luxury car salesman*) in isolation from other aspects of prenominal structure.

7 Prepositions

All five grammatical relations formats treat preposition phrases in pretty much the same way: (a) there is a GR link from the head of which the PP is a complement or modifier to the preposition itself (the HPSG representation has this link going in the opposite direction for PP modifiers, but the principle is the same); and (b) there is a link from the preposition to its complement NP. For example, the noun phrase *experts in Congress* is annotated as follows:

```
(8)  ncmod(experts,in)
     dobj(in,Congress)
```

The only PPs which have been handled differently are agentive *by*-PPs of passive participles, which are either normalised or treated using a special, construction-specific GR.

Note however that all prepositions are not equal when it comes down to representing the predicate-argument structure of a sentence. In a nutshell, some prepositions are predicators (e.g. *experts in Congress*) whereas others are simply syntactic role markers (e.g. *a workout of the Suns*). Ideally, we would want a GR system which marks this distinction, for example by annotating predicator prepositions as lexical heads and ignoring role-marking prepositions altogether. The only GR scheme which attempts to make this distinction is the PARC system, which has a `ptype` feature for every preposition with two possible values, `semantic` and `non-semantic`. However, this does not appear to have been annotated consistently in the PARC dataset — the only examples of non-semantic prepositions are agentive *by*-PPs of passive participles.

8 Conclusion

We have proposed a set of principles for developing a grammatical relation annotation system for use with both shallow and deep semantic interpretation systems, in particular a tutorial dialogue system. We then evaluated five different GR schemes from the dependency parsing literature based on how well they handle a number of 'deep' syntactic phenomena implied by these principles,

	noun to modifier			modifier to noun			predicative		
	A	V	N	A	V	N	A	V	N
RASP	ncmod			-	`ncsubj` etc	-	-	`ncsubj` etc	-
HPSG	-			`a_arg1`	`v_arg1` etc	`n_arg1`	`a_arg1`	`v_arg1` etc	`n_arg1`
CCG	`N/N`	-	`N/N`	-	`S\NP` etc	-	`Sadj\NP`	`S\NP` etc	?
PARC	`adjunct`		`mod`	-			`subj`	`subj`	?
Stanf	`amod`		`nn`	-			`nsubj`	`nsubj`	?

Table 3: Representation of prenominal modifiers

i.e. passive, control and raising, noun modification, and meaningful vs. non-meaningful prepositions. We conclude that none of the proposed GR annotation schemes contains everything we require for deep semantic processing, although each of the features/distinctions we included in our list of desiderata is provided by at least one system.

Many of the deep syntactic phenomena discussed here are known issues for shallow semantic tasks like semantic role labelling. For example, passive constructions are a recognised source of noise in semantic role labelling systems (Gildea and Jurafsky, 2002), and resolving controlled subjects provides more data for training models of selectional restrictions, which are known to be useful features for role labelling. More generally, Chen and Rambow (2003) demonstrate that a focus on 'deep' syntactic features results in a more accurate stochastic semantic role labeller than using surface information alone.

Note also that the deep grammatical role representation proposed here is meant to be 'theory-neutral', in the sense that it was not influenced by any one of the competing grammar formalisms to the exclusion of the others. Indeed, it should be a straightforward task to write a grammar using either the HPSG, LFG, CCG or RASP-style underlying formalism which can produce an output representation consisting of deep relations, constructed in a purely compositional manner. Indeed, the syntactic phenomena discussed in this paper are those which form the basis of numerous introductory textbooks on English generative syntax (Haegeman, 1994; Sag and Wasow, 1999; Bresnan, 2000). In addition, the phenomena which form the basis of the analysis in this paper were among those which had been the focus of a significant amount of attention in the development of the semantic interpretation system underlying our domain-independent tutorial dialogue system. Other issues which were considered, but for which we lack space to discuss in detail include: (a) expletive pronouns should be ignored, i.e. the subject pronouns in 'impersonal' verb constructions like *It is raining* or *It's great that John loves Mary* should not be seen as the target of deep grammatical relations; (b) unbounded dependencies should be resolved, i.e. in the relative clause *the woman Bill thinks John loves* there should be an object relation between the embedded verb *loves* and its extracted object *woman*; (c) restrictive and non-restrictive modification (including apposition) should be distinguished, since the latter is not relevant for reference resolution; and (d) certain subsentential conjunctions need to be compiled out (for examples like *electronic, computer and building products*).

Finally, we recognise that, in many cases, it is possible to transform parser representations into our desired format. For example, if the parser output tells us that a given verb form is a passive participle, we can use this information to remap the surface relations, thus retrieving the underlying predicate-argument structure. However, we prefer a system where this kind of post-processing is not needed. Reasons for this include the increased potential for error in a system relying on post-processing rules, as well as the need to have both detailed documentation for how each parser output format handles particular constructions, as well as a comprehensive mapping schema between representations. Having a community standard for GR-based parser output is an essential element of future parsing technology, and to be practically useful in a range of semantic interpretation tasks, this standard should involve 'deep' syntactic distinctions of the kind discussed in this paper.

9 Acknowledgements

References

Allen, James, Myroslava Dzikovska, Mehdi Manshadi, and Mary Swift. 2007. Deep linguistic processing for spoken dialogue systems. In *Proceedings of the ACL'07 Workshop on Deep Linguistic Processing*.

Bos, Johan and Tetsushi Oka. 2002. An inference-based approach to dialogue system design. In *Proceedings of COLING'02*.

Bresnan, Joan. 2000. *Lexical-Functional Syntax*. Basil Blackwell.

Briscoe, Ted, John Carroll, and Rebecca Watson. 2006. The second release of the RASP system. In *Proceedings of the COLING/ACL'06 Interactive Presentation Sessions*.

Callaway, Charles B., Myroslava Dzikovska, Elaine Farrow, Manuel Marques-Pita, Colin Matheson, and Johanna D. Moore. 2007. The Beetle and BeeDiff tutoring systems. In *Proceedings of SLaTE'07*.

Chang, N., J. Feldman, R. Porzel, and K. Sanders. 2002. Scaling cognitive linguistics: Formalisms for language understanding. In *Proceedings of ScaNaLU'02*.

Chen, John and Owen Rambow. 2003. Use of deep linguistic features for the recognition and labeling of semantic arguments. In *Proceedings of EMNLP'03*.

Copestake, Ann, Dan Flickinger, Carl Pollard, and Ivan A. Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language and Computation*, 3:281–332.

de Marneffe, Marie-Catherine, Bill MacCartney, and Christopher D. Manning. 2006. Generating typed dependency parses from phrase structure parses. In *Proceedings of LREC'06*.

Dzikovska, Myroslava O., Gwendolyn E. Campbell, Charles B. Callaway, Natalie B. Steinhauser, Elaine Farrow, Johanna D. Moore, Leslie A. Butler, and Colin Matheson. 2008. Diagnosing natural language answers to support adaptive tutoring. In *Proceedings of FLAIRS'08 special track on Intelligent Tutoring Systems*.

Erk, Katrin and Sebastian Padó. 2006. SHAL-MANESER - a toolchain for shallow semantic parsing. In *Proceedings of LREC'06*.

Gildea, Daniel and Daniel Jurafsky. 2002. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3).

Haegeman, Liliane. 1994. *Introduction to Government and Binding Theory*. Basil Blackwell, 2nd edition edition.

Hockenmaier, Julia and Mark Steedman. 2007. CCG-bank: a corpus of CCG derivations and dependency structures extracted from the Penn Treebank. *Computational Linguistics*, 33(3).

Jordan, Pamela, Maxim Makatchev, Umarani Pappuswamy, Kurt VanLehn, and Patricia Albacete. 2006. A natural language tutorial dialogue system for physics. In *Proceedings of FLAIRS'06*.

Kaisser, Michael and Bonnie Webber. 2007. Question answering based on semantic roles. In *Proceedings of the ACL'07 Workshop on Deep Linguistic Processing*.

Kay, Martin, Jean Mark Gawron, and Peter Norvig. 1994. *Verbmobil: A Translation System for Face-To-Face Dialog*. CSLI Press, Stanford, CA.

King, Tracy Holloway, Richard Crouch, Stefan Riezler, Mary Dalrymple, and Ronald M. Kaplan. 2003. The PARC 700 dependency bank. In *Proceedings of EACL'03*.

Miyao, Yusuke. 2006. *From Linguistic Theory to Syntactic Analysis: Corpus-Oriented Grammar Development and Feature Forest Model*. Ph.D. thesis, University of Tokyo.

Rosé, C. P., D. Bhembe, S. Siler, R. Srivastava, and K. VanLehn. 2003. The role of why questions in effective human tutoring. In *Proceedings of AIED'03*.

Sag, Ivan A. and Thomas Wasow. 1999. *Syntactic Theory: A Formal Introduction*. CSLI.

Seneff, Stephanie. 1992. TINA: A natural language system for spoken language applications. *Computational Linguistics*, 18(1).

Surdeanu, Mihai, Sanda M. Harabagiu, John Williams, and Paul Aarseth. 2003. Using predicate-argument structures for information extraction. In *Proceedings of ACL'03*.

Wolska, Magdalena and Ivana Kruijff-Korbayová. 2003. Issues in the interpretation of input in mathematical dialogs. In Duchier, Denys, editor, *Prospects and advances in the syntax/semantics interface. Lorraine-Saarland Workshop Series proceedings*.

Wolska, Magdalena and Ivana Kruijff-Korbayová. 2004. Analysis of mixed natural and symbolic language input in mathematical dialogs. In *Proceedings of ACL'04*.

Yeh, Peter Z., Bruce Porter, and Ken Barker. 2005. Matching utterances to rich knowledge structures to acquire a model of the speaker's goal. In *Proceedings of K-CAP'05*.