Underspecification of intersective modifier attachment: Some arguments from German

Berthold Crysmann
DFKI, Saarland University

Proceedings of the 11th International Conference on
Head-Driven Phrase Structure Grammar

Center for Computational Linguistics, Katholieke Universiteit Leuven

Stefan Müller (Editor)

2004

Stanford, CA: CSLI Publications

pages 378–392

Crysmann, Berthold. 2004. Underspecification of intersective modifier attachment: Some arguments from German. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 378–392. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2004.21.

Abstract

In this paper, I shall discuss the semantic attachment of intersective modifiers in German coherent constructions. I shall show that a purely syntactic solution to the observable attachment ambiguity is undesirable for reasons of processing efficiency and/or massive spurious ambiguity. Instead, I shall follow Egg and Lebeth (1995) and propose an extension to Minimal Recursion Semantics, permitting the expression of underspecified semantic attachment. This rather trivial move, as we shall see, will not only be preferable for processing reasons, but it will also be more in line with the spirit of underspecified semantics, effectively providing a compact representation of purely semantic distinctions, instead of unfolding these distinctions into a rain forest of tree representations and derivations. I will present an implementation of the underspecification approach integrated into the German HPSG developed at DFKI and compare its efficiency to an alternative implementation where semantic attachment is unfolded by means of retrieval rules.

1 Intersective modifiers and word order

It is a well-known property of German that order in the Mittelfeld is extremely free: although some restrictions do seem to exist as to the relative order in which a verb's complements can appear, it is by now generally accepted that the linearisation constraints regulating order within the Mittelfeld should best be conceived of as soft constraints or performance preferences (Uszkoreit, 1987). The word-order freeness of German is further multiplied by the fact that auxiliaries, modals, control and raising verbs may or must construct coherently (Kiss, 1994, 1995, Müller, 1999, Hinrichs and Nakazawa, 1990), a construction that is modelled by means of Hinrichs/Nakazawa-style argument composition. What is more, inherent and inherited arguments can, again, undergo scrambling, thus, in principle, arguments of the upper and lower verbs may appear in any order.

One of the fundamental empirical tests for the coherent construction — besides scrambling of arguments, of course — builds on the interpretation of modifiers. With a few exceptions, e.g. the marker of sentential negation *nicht* 'not', there does not appear to be any general positional restriction on the distribution of modifiers in the Mittelfeld: as a rule of thumb, modifiers can appear just about anywhere between the left and right sentence bracket, demarcated by a complementiser or a finite verb and the sentence-final verb cluster. Independently of position, however, modifiers in the coherent construction often display a systematic ambiguity between high and low attachment.

[†]This work has been carried out as part of the DFKI project QUETAL, funded by the German Federal Ministry for Education and Research (BMBF), and the EU-project DEEPTHOUGHT at the Department for Computational Linguistics, Saarland University. The ideas presented in this paper have benefited a lot from discussion with various people. I would therefore like to thank Ann Copestake, Markus Egg, Dan Flickinger, Stefan Müller, Stephan Oepen and Emily Bender for their comments and suggestions.

(1) Peter hat es im Labor blitzen sehen
Peter has it in.the lab flash saw
'Peter saw some flashes/lightning in the lab' (P"utz, 1982, 340)

As exemplified by the datum above, the PP *im Labor* can modify either the seeing event, or the flashing event: under the first interpretation, Peter is in the lab observing some lightning or flashes somewhere else (possibly outside), whereas under the latter, the flashes are in the lab, with the locus of the observer unspecified.

Although, in (1) the modifier is adjacent to the verb cluster, permitting us to model the two semantic interpretations by means of high or low syntactic attachment, this is not always the case: as illustrated in (2), a flipped auxiliary may intervene (Hinrichs and Nakazawa, 1994, Kathol, 2000, Meurers, 2001), making adjunction to the most deeply embedded cluster element impossible.[1] Still, the modifier displays the same semantic attachment ambiguity as in the example in (1) above.

(2) weil Peter es im Labor [hat [[blitzen] sehen]]
because Peter it in.the lab has flash saw
'because Peter saw some flashes/lightning in the lab'

The very same can be observed with scrambling in the Mittelfeld:

(3) a. weil Peter im Labor es blitzen sah
because Peter in.the lab it flash saw
'because Peter saw some flashes/lightning in the lab'

b. weil im Labor Peter es blitzen sah
because in.the lab Peter it flash saw
'because Peter saw some flashes/lightning in the lab'

Independent of surface position, and, therefore, constituency, modification of upstairs and downstairs verb is equally possible.

Similar evidence against a purely syntactic approach to intersective modifier attachment is provided by Egg and Lebeth (1995):

(4) Sollen wir im März noch einen Termin machen?
shall we in March an appointment make
'Should we schedule a meeting in March?' (Egg and Lebeth, 1995)

The sentence in (4) is three-ways ambiguous: the PP adjunct *im März* 'in March' may modify the appointment (*Termin*), the scheduling event (*ausmachen*), or even the modal (*sollen*). Under standard assumptions of phrase structure in

[1] For the purposes of this paper I will concentrate foremost on versions of HPSG without word order domains. As far as I can tell, the issues raised within the scope of this paper are by-and-large the same for linearisation approaches and true movement analyses (see below).

the German modal constructions (Kiss, 1994), only attachment to *sollen* should be available. Attachment to the main verb infinitive, however, can only be derived by making otherwise unmotivated assumptions about phrase structure, namely that modals optionally take a VP constituent as their complement.

It should be clear that the data presented thus far constitute a syntax-semantics mismatch: ceteris paribus, modification of the downstairs verb obviously conflicts with straightforward rule-by-rule compositionality. Thus, some more elaborated mechanisms are called for to derive the full set of interpretations, independent of constituency in the Mittelfeld.

1.1 Storage and Retrieval

One such extension has been proposed in Kiss (1995): to overcome the kind of problem just sketched, he proposes to collect modification targets in a special storage feature from which they can be retrieved whenever a modifier is attached in syntax. Introduction of modification targets onto the storage works in tandem with verb complex formation. Though certainly a viable solution at the time, nowadays, such an approach is not anymore fully attractive, with the Cooper-storage being successfully supplanted by much more concise underspecified descriptions. Furthermore, in a computational setting,[2] retrieval during parsing can be quite costly, as the exact number of modification targets is locally not always known in bottom-up parsing. Owing to the fact that entire verb clusters can be extracted into the Vorfeld, the complexity of the extracted cluster is unknown at the point where the Mittelfeld is constructed. Thus, whenever Partial VP Fronting can be hypothesised during parsing, the number of available modification targets to be assumed locally will be equal to the maximum complexity of verbal clusters in German.

(5) Blitzen sehen hat Peter es im Labor.
flash see has Peter it in.the lab
'Peter saw some flashes/lightning in the lab'

Even if we can put an upper bound on verb cluster complexity — the most complex cluster I found in Meurers (1997) consisted of 5 elements in total —, it should be kept in mind that retrieval of modification targets during parsing will increase by this factor not only the number of head-modifier edges themselves but also the number of chart items that can be transitively derived from these edges. Although the overall frequency of partial VP fronting in German is not that high, local ambiguity is unaffected by this, due to the unbounded nature of the process: even the chart of "harmless" sentences without any PVP fronting is characterised by an incommensurate number of PVP hypotheses.

[2] As a point of reference I use the fastest processing platform for HPSG grammars currently available, namely PET (Callmeier, 2000) together with the development platform LKB (Copestake, 2001). As for the grammar, I will assume the large-scale grammar of German, developed in the Verbmobil context by Müller and Kasper (2000), which has been ported to LKB/PET by Stefan Müller and subsequently enhanced by Berthold Crysmann (Crysmann, 2003, to appear).

Based on these two objections, we can discard the storage-retrieval approach as a suboptimal solution, at least, unless a more efficient and elegant solution can be found.

1.2 Scrambling as movement

Another obvious way to attack the issue is to analyse scrambling as movement, akin to analyses carried out in the generative paradigm. Besides the issue whether or not one should treat essentially local order phenomena on a par with unbounded dependencies, an extraction-based approach will introduce a fair amount of spurious ambiguity into the grammar: unless we can canonicalise the introduction of modifier gaps in a highly restrictive fashion, regulating the relative attachment for every pair of different modifiers, even simple sentences with only one modification target but two intersective modifiers will end up with two syntactically different, yet semantically identical analyses. Worse, the amount of spurious ambiguity thus introduced will be factorial to the number of modifiers present. Finally, in standard bottom-up parsing, the number of modifier gaps to be introduced cannot be known a priori, so we either have to artificially limit the number of scrambled modifiers, or else suffer a termination problem.[3] Thus, we can safely discard this latter type of analysis altogether.

2 Modifier interaction

Having established so far that neither a Kiss-style storage and retrieval mechanism nor a movement-based analysis can qualify as optimal solutions to the empirical problem, I will now move on and explore, if and how the treatment of intersective modifier attachment can be likened to that of quantifiers and scopal modifiers, ultimately leading towards a treatment in terms of underspecification.

An important question to be addressed in this context is whether high vs. low attachment of a modifier interacts with the attachment of other modifiers in the sentential domain, or whether different modifiers rather enjoy the same range of attachment possibilities independently of each other.

At least for the interpretation of scopal modifiers, it has repeatedly been claimed (Müller, 1999, Kasper, 1994, Müller, 2004) that scope in the German Mittelfeld is determined from left to right. Although I do not doubt that this is the case more often than not, counter-examples to this allegedly hard constraint of German syntax can easily be provided:[4]

[3] This is at least true in formalisms without lazy evaluation, such as LKB and PET.

[4] The data presented in this section have each been confirmed by 4 native speakers, in addition to my own intuitions. 3 of these subjects are non-linguists, the other does not actively work on German syntax.

(6) Da muß es schon erhebliche Probleme mit der Ausrüstung gegeben
there must it severe problems with the equipment given
haben, da wegen schlechten Wetters ein Reinhold Messner niemals
have since because.of bad weather a Reinhold Messner never
aufgäbe.
give-up.would

'There must have been severe problems with the equipment, since someone like Reinhold Messner would never give up just because of the bad weather.'

Example (6) contains two scopal modifiers, *niemals* 'never' and the *wegen*-PP, a causative operator. Although linear order would suggest interpretation of *niemals* in the scope of the *wegen*-PP, the preferred reading, however, has the relative scope of these two modifiers reversed.

Similar evidence can be found concerning the interaction of intersective modifiers and scopal modifiers.

(7) Stefan ist wohl deshalb krank geworden, weil er äußerst
Stefan is presumably therefore ill become because he extremely
hart wegen der Konferenz in Bremen gearbeitet hat.
hard because.of the conference in Bremen worked has

'Stefan probably only became ill, because he worked extremely hard because of the conference in Bremen.'

As already observed by Kasper (1994, p. 47), embedding of a causal modifier under an intersective manner adverb is ruled out for semantic reasons. Still, a sentence like the one in (7) is completely well-formed, the only available interpretation having the intersective modifier within the scope of the causal modifier.

Thus, we can conclude that the left-to-right scope rule is but a performance preference, however strong.

If we return now to the issue of intersective modifier attachment, we find that here again, the left-to-right rule does not always restrict the range of possible attachments.

(8) Bei dem Wetter wird ohne Regenmantel ein besorgter Vater seine
with the weather will without macintosh a caring father his
Kinder niemals aus dem Haus gehen lassen.
children never out the house go let

'In this weather, a caring father will never let his children go out without a macintosh.'

As illustrated by the example in (8), the PP-modifier *ohne Regenmantel* 'without a macintosh' can and, given world knowledge, must attach to the lower verb *ausgehen* 'go out', despite the intervention of the scopal temporal modifier *niemals*

'never', which scopes over the entire verbal complex, at least under the highly preferred reading. It appears, thus, that the interpretation of an intervening modifier does not interfere with the availability of the downstairs verb as a modification target.

(9) Der diensthabende Beamte gab zu Protokoll, daß in der
the policemen who was on duty gave to protocol that in the
Dachwohnung zum fraglichen Zeitpunkt ein Rentner von der
loft at.the time in question a retired man from the
anderen Straßenseite aus die Angeklagte mehrmals auf das Opfer
opposite side of the road the accused repeatedly on the victim
einstechen sah.
stab saw

'The policemen who was on duty noted that a retired man witnessed from the opposite side of the road that, in the apartment under the roof, at the time in question, the accused stabbed the victim several times.'

This last finding can be replicated with intersective modifiers as well. Here, the PP *von der anderen Straßenseite aus* 'from the opposite side of the road' must modify the seeing event[5] The locative PP *in der Dachwohnung*, however, under the most preferred interpretation, attaches semantically to the downstairs stabbing event.

As we will see below, the observable independence of multiple modifiers with respect to the availability of modification targets will be highly advantageous in the context of an underspecification approach.[6]

3 A proposal

Within current MRS (Copestake et al., 1998, 2001, to appear), the treatment of intersective modifiers essentially assumes that syntactic and semantic structure be homomorphic, as far as attachment is concerned. Thus, at present, this semantic description language does not provide any tools out of the box to address the issues raised in this paper. Essentially, an intersective modifier has its unified with that of the head daughter, whereas its feature is unified with the head daughter's . Both these links are hard-wired. Although this assumption works quite well for highly configurational languages such as English, a treatment along these lines is actually not too well-equipped to deal with non-configurational languages, such as German, where a syntactic solution, as detailed above, will be both inefficient and inelegant.

[5]Unless we want to make the unlikely assumption that the accused has arms as long as Mr Tickle's.

[6]Conversely, this independence is quite disadvantageous to a storage and retieval approach, since the number of targets to be considered will not shrink.

If we want to provide an underspecified representation of intersective modifier attachment as well, all we need to do is define a data structure suitable for distributing a modifier's and values over the set of / pairs contributed by the modification targets. These / pairs are best encapsulated as a data structure of their own, which I will call *anc(hor)*, following a proposal of Kiss (in press, 2003) for a semantic treatment of relative clause extraposition. A natural place to represent constraints on the possible attachment of an intersective modifier is the - feature, currently hosting *qeq* constraints only.

(10)
$$\begin{bmatrix} anc & \\ & handle \\ & index \end{bmatrix}$$

(11)
$$\begin{bmatrix} isect\text{-}mod & \\ \text{-} & anc \\ \text{-} & list\left(anc\right) \end{bmatrix}$$

With a basic underspecified representation in place, all we have to do in syntax is to define the list of target anchors, and introduce an appropriate constraint into - , whenever an intersective modifier is syntactically attached.

An implementation of these two steps is quite straightforward. Let us begin with the definition of the list of target anchors: in order to avoid traversal of the - list, I will invoke an auxiliary feature _ - , which I will assume to be located under | . The value of this feature is, again, a list of *anc*. Verbs that do not construct coherently, or, that do not take any verbal complements at all, will have a singleton _ - list, where the and features of the only *anc* are unified with the and of the verb itself. Verbs, however, that do construct coherently, will specify an open list, where the first element is again linked to the verb's own and . Yet, the rest of the list will be structure-shared with the | |_ - of the verbal complement, represented under (see, e.g., Müller, 1999, Kathol, 1999, for motivation of this valence feature).

Now, whenever an intersective modifier gets syntactically attached, we simply add a new *isect-mod* constraint to the - list, unifying its - | and - | with the and of the modifier's . The - feature of this constraint will just be structure-shared with the _ - feature of the head-daughter's .

As to scopal modifiers, nothing fancy has to be done here: all one needs to do is to have the scopal modifiers outscope the lowest handle in the verb cluster. As a consequence, scopal modifiers will be able to assume any intermediate scope. The attentive reader will have noticed that the solution proposed here implicitly assumes independent phrase structure schemata for intersective and scopal modification. However, this is not really new: as argued by Copestake et al. (to appear), such a move is independently required to ensure a sound treatment of Kasper's problem.

Comparing the approach advanced here with the alternative solutions refuted above, we find that it is highly similar to a storage approach, with the retrieval step delayed to post-syntactic semantic resolution. Under a processing perspective, however, such a move is highly advantageous: first, purely semantic distinctions are effectively encapsulated in the -value, rather than expanded into different tree representations or derivations. Furthermore, as we have seen above, proper attachment can often not be resolved on the basis of sortal restrictions. Rather, it is world or discourse knowledge that decides on the most likely attachment. Second, unfolding the set of possible attachment as part of MRS resolution will be much more efficient, as the issue of local ambiguity and the adverse effects on the search space encountered in parsing are simply non-issues: the size of the resulting parse forest and the associated MRS substructures are actually tiny in comparison to the chart that needs to be explored to deliver them.

The approach to underspecified intersective modifier attachment suggested here bears the further potential of reducing processing cost in the context of relative clause extraposition. As detailed in Crysmann (to appear), a considerable amount of the additional cost required by my implementation of a Kiss-style approach to German relative clause extraposition is spent on the retrieval of suitable anchors during parsing: with an upper limit of the 5 most recent anchors introduced, the relative cost of integrating this construction is reflected in an increase by of executed chart items by around 12.7% on the Babel test suite. [7] With the kind of underspecification advocated here, retrieval can, again, be postponed to a semantic post-processing step, avoiding retrieval costs and ensuring termination without having to impose arbitrary limits. It is of note, in this context, that the data structure suggested here for the expression of underspecified modifier attachment, i.e. - / pairs, also plays a crucial role under a semantic approach to relative clause extraposition, like the one suggested by Kiss (in press, 2003).

4 Evaluation

In the preceding sections, we have seen that the semantic attachment of intersective modifiers cannot be derived on the basis of surface phrase structure alone. We have discussed three alternative approaches — movement, storage and retrieval, and underspecification — and concluded that among these three, the movement-based approach should be rejected a priori, since it will lead to massive spurious ambiguity, as well as suffer from termination problems on systems without lazy evaluation, such as LKB or PET. This leaves us with only two options: a Kiss-style storage and retrieval approach, and underspecified attachment.

In order to substantiate the expected performance gains of the underspecification approach over a storage and retrieval mechanism (cf. Egg and Lebeth, 1995),

[7] It should be kept in mind that the impact of retrieval rules on relative clause extraposition is of a must lesser degree than what we can expect for ordinary modifiers, owing to the fact that these rules are only applied once a rather large sentential constituent has already been built.

LKB/PET implementations of both variants have been provided and subsequently evaluated. As evaluation corpora, I have chosen the manually constructed Babel test-suite (Müller, 2004; 758 test items), as well as a subset of the Verbmobil spoken language corpus (VM CD-15; 2233 test items).

The underspecification variant is compared to a baseline where semantic attachment is isomorphic to syntactic attachment, enabling us to estimate the minimal cost associated with a sound treatment of intersective modifier attachment.

All test runs have been performed on an Intel Pentium 4 M with 1 GB RAM, running Linux 2.4.26. The version of PET used for the tests dates from July 2003. Test results have been collected and evaluated with [incr tsdb()] (Oepen and Flickinger, 1998).

4.1 Baseline vs. Underspecification

The implementation of underspecified modifier attachment was derived directly from a baseline implementation where modifiers could only semantically attach to the label and index of their syntactic sister (see Müller and Kasper (2000), Müller (2004), Crysmann (2003, to appear) for further details on the German HPSG developed at DFKI). Most obviously, the grammar was extended with a mechanism to collect lists of target anchors during verb cluster construction. Furthermore, the distinction between intersective and scopal modification, which was hitherto performed at the level of lexical types only, is now replicated at the level of phrase structure schemata. Recall that such a move is independently required to provide a solution to Kasper's problem along the lines proposed in Copestake et al. (to appear). While such a move is virtually cost-neutral for local head-adjunct structures — selection of intersective vs. scopal head-adjunct schemata is entirely determined by the lexical type of the modifier —, this is not the case with adjunct extraction: since the type of modifier is not known locally, we expect an increase in local ambiguity during parsing.[8] Finally, in order to control for spurious ambiguity, syntactic attachment to downstairs verbs in the cluster is blocked. Thus, even those attachment ambiguities that could in principle be resolved at the level of bare phrase structure are now taken care of by target anchor percolation and underspecified semantic attachment.

The results for these two grammars are given in tables 1 and 2. As depicted in table 1, both grammars have roughly the same coverage on the two corpora used for evaluation. While lexical ambiguity of the grammars is the same, syntactic ambiguity is slightly reduced under the underspecification approach. This slight reduction can be attributed to the elimination of syntactic attachment ambiguities for cluster-adjacent adjuncts.

With respect to performance[9], we record mild efficiency gains on both corpora.

[8]The additional cost associated with adjunct extraction could in principle be eliminated by moving the semantic effect of adjunction from the gap site to the filler site. Due to time constraints, however, this solution was not explored in the grammars presented here.

[9]Throughout this paper, tasks represent the average number of executed tasks per item, time is

Test suite	words	lex amb	Baseline cov	Baseline amb	Underspecification cov	Underspecification amb
VM 15	5.18	2.96	84.7	9.92	84.8	9.36
Babel	6.76	2.99	82.9	3.68	82.9	3.67

Table 1: Baseline vs. Underspecification: Coverage & Ambiguity

Test suite	items	Baseline tasks	Baseline time	Baseline space	Underspecification tasks	Underspecification time	Underspecification space	Factor tasks	Factor time	Factor space
VM 15	2233	12896	.83	24172	13650	.76	22969	1.058	.906	.95
Babel	758	4117	.21	6731	3639	.18	5755	.884	.849	.855

Table 2: Baseline vs. Underspecification:
Performance (including non-exhaustive parses)

Thus, we can conclude that the underspecified approach provides an efficient way to arrive at a complete representation of the attachment potential of intersective modifiers.

4.2 Underspecification vs. Storage & Retrieval

In a second implementation step, I have derived a third version of the grammar, where semantic attachment is fully resolved in syntax. Since percolation of modification targets is unaffected by the way that a modifier gets semantically bound to a target anchor, I have derived this version from the underspecification approach.

The only changes that needed to be performed involve the unfolding of syntactically underspecified attachment into distinct intersective head-adjunct rules: thus, instead of a single head-adjunct rule that simply inserts an appropriate modification constraints into the MRS, representing the distribution of the modifiers anchor over the list of target anchors, we now have to enumerate, by means of distinct syntactic rules, the range of possible semantic attachments. In order to ensure a fair comparison, I have limited access to percolated anchors to the first 3 target anchors. This number should correspond quite well to the maximum complexity of verb clusters observed in corpora, which is 4 (Müller, p.c.), considering that verb clusters of this size will most probably include tense or passive auxiliaries, which I assume to share their event variable with that of the main verb they combine with. In sum, the transition from a syntactically underspecified analysis of intersective modifier attachment to a storage and retrieval approach involves tripling the number of head-initial and head-final intersective head-adjunct rules, as well as that of intersective adjunct extraction rules. As a result, local ambiguity during parsing will increase by this figure, in the worst case.

The results of the comparison between the underspecification and storage/retrieval approaches are summarised in tables 3–5.

average parse time per item in seconds, and space indicates the average space consumption per item in kB.

As shown in table 3, overall coverage decreases with the storage and retrieval approach. This result is directly related to the decrease in parsing efficiency to be described below, reflecting the exhaustion of available resources (70,000 passive edges) before a result could be delivered.

			Underspecification		Storage/Retrieval	
Test suite	words	lex amb	cov	amb	cov	amb
VM 15	5.18	2.96	84.8	9.37	83.6	11.48
Babel	6.76	2.99	82.9	3.68	82.5	3.48

Table 3: Underspecification vs. Storage & Retrieval: Coverage & Ambiguity

	items	Underspecification			Storage/Retrieval			Factor		
Test suite		tasks	time	space	tasks	time	space	tasks	time	space
VM 15	2233	13650	.76	22969	24892	1.24	32324	1.824	1.644	1.407
Babel	758	3639	.18	5755	5942	.27	7870	1.633	1.535	1.368

Table 4: Underspecification vs. Storage & Retrieval:
Performance (including non-exhaustive parses)

With respect to performance, a direct comparison on all test items reveals that the syntactic retrieval of anchors is quite costly, leading to an increase by more than 50% in parse time. However, as we have already observed above, the difference in coverage the two grammars display on Verbmobil data is due to the fact that the less efficient storage and retrieval approach reaches the upper limit of 70,000 passive edges much more often than the grammar implementing the underspecification approach. It appears thus, that the less efficient grammar might benefit from a ceiling effect here, since, on items where the available resources are exhausted, this grammar cannot possibly get worse than the time it takes to build up 70,000 passive edges.

In order to get a more accurate picture of the relative performance of the two grammars, I therefore provide additional performance data, derived from those test items where both grammars had been able to explore the entire parse space within the given limit.

	items	Underspecification			Storage/Retrieval			Factor		
Test suite		tasks	time	space	tasks	time	space	tasks	time	space
VM 15	2070	3330	.17	5456	9277	.43	12069	2.786	2.594	2.212
Babel	758	2981	.14	4680	4791	.22	6203	1.607	1.508	1.325

Table 5: Underspecification vs. Storage & Retrieval:
Performance (intersection of exhaustively parsed items)

The intersection of sentences exhaustively parsed by both grammars provides a more reliable comparison, showing that, on Verbmobil data of moderate complexity, the underspecification outperforms the storage and retrieval approach to intersective modifier attachment by a factor of 2.6.

5 Conclusion

In this paper I have argued for an extension to Minimal Recursion Semantics (MRS; Copestake et al., 1998, 2001, to appear) permitting the expression of underspecified intersective modifier attachment similar to the proposal of Egg and Lebeth (1995). I have argued on the basis of German modifiers in the coherent construction that a complete, compact, and efficiently processable solution to the attachment ambiguity problem necessitates a treatment in underspecified terms. Furthermore, we have seen that the attachment potential of each individual modifier in the sentential domain is independent of the other modifiers in this domain. This particular finding has paved the way for a straightforward analysis in terms of MRS, enhanced by a new type of handle constraint, recording a modifiers anchor, together with the target anchors it can distribute over. This proposal, in contrast to the alternative syntactic solutions discussed in the text, puts the treatment of the issue much more in line with the spirit of underspecified semantics, namely to provide a compact representation of entirely semantic distinctions.

The proposal has been implemented as part of the German HPSG developed at DFKI, and systematically compared to an alternative approach, involving syntactically resolved attachment to percolated anchors. As detailed by the evaluation results presented in this paper, the underspecification approach outperforms the alternative syntactic solution by a factor between 1.5 and 2.6.

I have further argued that this particular proposal can also be put to use in the context of German relative clause extraposition. Furthermore, I conjecture that such an approach can be fruitfully applied to any other language featuring complex predicate formation of the argument composition type.

References

Callmeier, Ulrich. 2000. PET — A Platform for Experimentation with Efficient HPSG Processing Techniques. *Journal of Natural Language Engineering* 6(1), 99–108.

Copestake, Ann. 2001. *Implementing Typed Feature Structure Grammars*. Stanford: CSLI Publications.

Copestake, Ann, Flickinger, Dan, Pollard, Carl and Sag, Ivan. to appear. Minimal Recursion Semantics: an introduction. *L & C* .

Copestake, Ann, Flickinger, Dan and Sag, Ivan. 1998. Minimal Recursion Semantics. An Introduction, proceedings of ESSLLI 10, Saarbrücken.

Copestake, Ann, Lascarides, Alex and Flickinger, Dan. 2001. An Algebra for Semantic Construction in Constraint-based Grammars. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics, Toulouse, France*.

Crysmann, Berthold. 2003. On the Efficient Implementation of German Verb Placement in HPSG. In *Proceedings of RANLP 2003*, pages 112–116, Borovets, Bulgaria.

Crysmann, Berthold. to appear. Relative Clause Extraposition in German: An Efficient and Portable Implementation. *Research on Language and Computation* .

Egg, Markus and Lebeth, Kai. 1995. Semantic Underspecification and Modifier Attachment Ambiguities. In J. Kilbury and R. Wiese (eds.), *Integrative Ansätze in der Computerlinguistik (DGfS/CL '95), Düsseldorf, Seminar für Allgemeine Sprachwissenschaft*, pages 19–24.

Hinrichs, E. and Nakazawa, T. 1990. Subcategorization and VP Structure in German. In Hughes, Shaun and Salmons (eds.), *Proceedings of the Third Symposium on Germanic Linguistics*, Amsterdam: Benjamins.

Hinrichs, Erhard and Nakazawa, Tsuneko. 1994. Linearizing AUXs in German Verbal Complexes. In John Nerbonne, Klaus Netter and Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, Lecture Notes, No. 46, pages 11–37, Stanford University: CSLI Publications.

Kasper, Robert T. 1994. Adjuncts in the Mittelfeld. In John Nerbonne, Klaus Netter and Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, Lecture Notes, No. 46, pages 39–69, Stanford University: CSLI Publications.

Kathol, Andreas. 1999. Linearization versus Phrase Structure in German Coordinate Constructions. *Cognitive Linguistics* pages 303–342.

Kathol, Andreas. 2000. *Linear Syntax*. Oxford: Oxford University Press.

Kiss, Tibor. 1994. Obligatory Coherence: The Structure of German Modal Verb Constructions. In John Nerbonne, Klaus Netter and Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, Lecture Notes, No. 46, pages 71–107, Stanford University: CSLI Publications.

Kiss, Tibor. 1995. *Infinite Komplementation. Neue Studien zum deutschen Verbum infinitum*. Linguistische Arbeiten, Tübingen: Niemeyer.

Kiss, Tibor. 2003. Phrasal Typology and the Interaction of Topicalization, Wh-movement, and Extraposition. In Jong-Bok Kim and Stephen Wechsler (eds.), *Proceedings of the 9th International Conference on Head-driven Phrase Structure Grammar, Kyung Hee University, Seoul, 5–7 August, 2002*, Stanford: CSLI Publications.

Kiss, Tibor. in press. Semantic Constraints on Relative Clause Extraposition. *Natural Language and Linguistic Theory* forthcoming.

Meurers, Detmar. 2001. To flip or not to flip: On the nature of irregularities in the German verbal complex. In *Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, CSLI Online Proceedings, pages 235–246, Stanford: CSLI Publications.

Meurers, Walt Detmar. 1997. Statusrektion und Wortstellung in kohärenten Infinitkonstruktionen des Deutschen. In Erhard Hinrichs, Walt Detmar Meurers, Frank Richter, Manfred Sailer and Heike Winhart (eds.), *Ein HPSG-Fragment des Deutschen. Teil 1: Theorie*, volume 95 of *Arbeitspapiere des SFB 340*, pages 189–248.

Müller, Stefan. 1999. *Deutsche Syntax — deklarativ*. Linguistische Arbeiten, Tübingen: Niemeyer.

Müller, Stefan. 2004. Continuous or Discontinuous Constituents? A Comparison between Syntactic Analyses for Constituent Order and Their Processing Systems. *Research on Language and Computation* 2(2), 209–257.

Müller, Stefan and Kasper, Walter. 2000. HPSG Analysis of German. In Wolfgang Wahlster (ed.), *Verbmobil: Foundations of Speech-to-Speech Translation*, pages 238–253, Berlin: Springer.

Oepen, Stephan and Flickinger, Dan. 1998. Towards Systematic Grammar Profiling. Test Suite Technology Ten Years After. *Journal of Computer Speech and Language* 12, 411–436.

P”utz, Herbert. 1982. Objektspr”adikate. In Werner Abraham (ed.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung*, pages 331–367, Tübingen: Gunter Narr Verlag.

Uszkoreit, Hans. 1987. *Word Order and Constituent Structure in German*. Stanford: CSLI Publications.