

Relative Clause Extraposition in German: an efficient and portable implementation

Berthold Crysmann

*German Research Center for Artificial Intelligence (DFKI) GmbH &
Computational Linguistics, Saarland University*

October 7, 2004

Abstract. In this paper, I propose an implementation of relative clause extraposition in German. The proposal builds on Kiss (in press) who treats relative clause extraposition as an anaphoric process by means of percolation of anchors to which the relative clause is bound. I discuss several sources of spurious ambiguity in Kiss's original formulation and suggest a two-step percolation of anchors that crucially distinguishes right-peripheral from central or left-peripheral percolation. Since extraposition is fairly productive, and phrase-structure alternates between head initial (prepositional phrases, V-initial) and head-final structures (postpositional phrases, V-final), German provides a good testing ground for techniques controlling spurious ambiguity that may easily be ported to languages where phrase structure is more canonical and/or extraposition more restricted. Finally, the performance of the Kiss-style approach is compared to an alternative implementation in terms of rightward movement, similar to Keller (1995).

Keywords: Extraposition, HPSG, grammar engineering

Extraposition of relative clauses is a discontinuous dependency that is quite well-attested cross-linguistically (see Kuno, 1974 for an overview). Although languages may differ as to the productivity of the process and the constraints they impose on the construction, the phenomenon can be observed in typologically diverse languages, such as English (Ross, 1967), German (Kathol, 2000; Kathol and Pollard, 1995; Keller, 1995; Müller, 1999), French, or Hungarian (Kiss, 1981; MacWhinney and Pleh, 1988). Thus, relative clause extraposition is found in highly configurational languages such as English, as well as in so-called free word order languages such as German or Hungarian. Furthermore, the phenomenon also appears to be independent of direction of headed-ness, featuring in canonically head-initial languages, as well as in languages where the position of the head is either variable or canonically final. As to the functional motivation for this type of discontinuous dependency, there is by now general agreement that avoidance of heavy center embedding is a major factor (MacWhinney and Pleh, 1988; Hawkins, 1994; Lewis, 1996; Gibson, 1998; Uszkoreit et al., 1998), accounting for a relatively high frequency of the construction in natural language corpora.¹

Despite the fact that relative clause extraposition is a typologically wide-spread phenomenon, it is quite surprising that, with the exception

of Stefan Müller’s Babel grammar (Müller, 2004), no large-scale HPSG grammar currently exists that takes this phenomenon into account. In this paper, I shall propose an efficient implementation of relative clause extraposition in a large-scale HPSG for German. A major concern for any computational treatment of extraposition phenomena is to control for spurious ambiguity: I shall argue that, owing to the variable position of heads in this language, the situations to be controlled for in German must be formulated independently of headedness. As a result, the principles of feature percolation developed on the basis of German are language-independent, thereby warranting a good degree of portability.

1. Relative clause extraposition without discontinuous constituents

The phenomenon of extraposition in German has been an area of central concern in the context of theoretical approaches to HPSG analysis of German. Although there is now some consent (Nerbonne, 1994; Kathol, 1995; Müller, 1999; Kathol, 2000) to regard extraposition of relative clauses as a word order phenomenon, these approaches all rely on the concept of word order domains, a mechanism that is so far absent from efficient implementations of the HPSG formalism, e.g. the runtime system PET (Callmeier, 2000) and the development platform LKB (Copestake, 2001), which both assume continuous constituents only.²

A viable, though not particularly efficient solution (see below) would be to simulate domain-based extraposition by way of rightward extraction (Keller, 1995): as relative clauses are always modifiers, and therefore optional, such a move will inevitably introduce a fair amount of local ambiguity. Since extraposition is not strictly local (Müller, 2004; Kiss, in press), but only upward bounded (Ross, 1967), a trace needs to be hypothesised for every NP contained in a clause. Owing to the nonlocality of the process, the ambiguity thus created can only be resolved at the clausal level.

An interesting theoretical alternative to linearisation-based or movement-based extraposition has recently been proposed by Kiss (in press; 2003).³ He argues that extraposition differs from extraction proper in that it does not seem to obey island constraints (1): Extraposition of relative clauses appears to be easily possible from constructions which otherwise place quite some strong restrictions on complement extraction, such as complex NPs (2a), or adjuncts (2b).⁴

- (1) a. Man hat [den Überbringer [der **Mitteilung**]]
 one has the messenger of.the message
 beschimpft, in **der** zur Räumung des
 insulted in which to.the evacuation of.the
 Parlaments aufgefördert wurde
 parliament demanded was
 ‘The person who delivered the message demanding the
 evacuation of parliament was insulted.’
- b. Hier habe ich [bei [den **Experimenten**]] faul auf
 here have I during the experiments lazily on
 der Wiese gelegen, bei **denen** die Schwerkraft
 the lawn laid at which the gravity
 überwunden wurde.
 overcome was
 ‘I was lying here lazily on the lawn during the experiments
 in which gravity was overcome.’
- (2) a. *Man hat [den Überbringer [der **Mitteilung**]]
 one has the messenger of.the message
 beschimpft, **daß** die Erde rund ist.
 insulted that the earth round is
 ‘The messenger was insulted who delivered the message
 that the world is a sphere.’ (Kiss, in press)
- b. *Hier habe ich [bei [den **Beobachtungen**]] faul auf
 here have I during the observations lazily on
 der Wiese gelegen, **daß** die Erde rund ist.
 the lawn laid that the earth round is
 ‘I was lying here lazily on the lawn during the observations
 that the world is a sphere.’

Instead he suggests modelling extraposition as an anaphoric process where a relative clause restricts an index contributed by an antecedent contained in the clause, an analysis that has also been suggested, for English, by Wittenburg (1987). This view is also supported by relative clause extraposition data involving split antecedents, as given in (3).

- (3) Ich habe heute einen **Mann** und gestern eine **Frau**
 I have today a man and yesterday a woman
 getroffen, **die** beide gerne Schach spielten.
 met who both happily chess played
 ‘I met a man today and a woman yesterday, who both liked
 playing chess.’

This basic idea is expressed by his principle of Generalised Modification:

(4) Generalized Modification:

The index of a modifying phrase has to be identified with a suitable index contained in the phrase to which the modifier is adjoined. (Kiss, in press)

Technically, this is achieved by means of percolating the `HANDEL` and `INDEX` of every NP in the clause via a set-valued nonlocal feature `ANCHORS` and requiring that the relative clause's `INDEX` and `HANDEL` features be token-identical with those of a member in `ANCHORS`.

(5) Anchor Projection Principle:

The anchors set of a headed phrase consists of the union of the anchors sets of the daughters less those anchors that are specified as `TO-BIND|ANCHORS` on the head-daughter. (Kiss, in press)

Upward boundedness, i.e., clause-locality of extraposition is modelled by imposing appropriate restrictions on phrasal and lexical signs capable of closing off a sentential projection. Thus, Kiss (in press) proposes that e.g. complementisers identify their `TO-BIND|ANCHORS` set with the `ANCHORS` set of their sentential complement, effectively blocking further percolation:

$$(6) \left[\begin{array}{l} \text{HEAD} \quad \text{comp} \\ \text{COMPS} \quad \left\langle \text{S} \left[\text{ANCHORS} \quad \boxed{1} \right] \right\rangle \\ \text{TO-BIND} \mid \text{ANCHORS} \quad \boxed{1} \end{array} \right]$$

A similar assumption is made for head-specifier structures, which he assumes license V2 structures.

Thus, sentences like (7) can effectively be ruled out, due to the lack of an appropriate anchor for the relative clause:

- (7) a. Ulrich hatte zugegeben, daß die **Karte** gestohlen war,
 Ulrich has admitted that the ticket stolen was
die er gefunden hatte, als er getrunken hatte.
 which he found had when he drunk has
 ‘While being drunk, Ulrich had admitted that the ticket
 he jhad found was stolen.’ (Kiss, in press)
- b. * Ulrich hatte zugegeben, daß die **Karte** gestohlen war,
 Ulrich has admitted that the ticket stolen was
 als er getrunken hatte, **die** er gefunden hatte.
 when he drunk has which he found had

(Kiss, in press)

The proposal advanced by Kiss (in press) is quite powerful, in that it is capable of modelling relative clause extraposition not only at the sentential level (8), but also at the level of PPs and participial clauses (see (9)).

- (8) Man hatte [**den Boten**] beschimpft, **der** den Befehl
 one has the messenger insulted who the command
 überbrachte.
 delivered
 ‘The messenger was insulted who delivered the command.’
 (Kiss, in press)
- (9) Man muß nicht gleich [[**eines Sachverständigen**]
 one must not immediately an expert
 wegen]_{PP}, **der** seinen Mund nicht halten kann, die gesamte
 because who his mouth not keep can the whole
 Konkurrenz kopfscheu machen.
 competitors nervous make
 ‘It does not make sense to confuse one’s competitors just
 because an expert could not keep quiet.’ (Kiss, in press)

Furthermore, from the point of view of NLP, the anaphoric perspective on relative clauses is to be preferred over any movement approach, as indices can be added without creating any local ambiguity, and processing cost is thus delayed until the point where a relative clause is actually found and an index needs to be retrieved.

2. Controlling Spurious ambiguity

A drawback of Kiss (in press), however, is that this approach introduces a fair amount of spurious ambiguity, a problem already recognised, in part, by the author. Spurious ambiguity arises from the fact that indices can, in principle, be retrieved at any point along the percolation path, giving additional extraposed analyses for modifiers that may equally be attached locally. Upon further scrutiny, we can actually distinguish two general cases here: one with head-final and one with head-initial structures.

- (10) Sie hatte den Termin verschoben am Montag.
 she had the appointment rescheduled on Monday
 ‘She had rescheduled the appointment on Monday.’ (Kiss, in press)

Under Kiss's (in press) approach, the above sentence receives two syntactic analyses, one with low attachment for the PP, corresponding to modification of the rescheduling event, and another one, with high attachment, capable of deriving modification of both the noun *Termin* and the verb *verschoben*.

A similar case of spurious ambiguity can be found with NPs. Here, low attachment (to N or N') is possible, on a par with modification of the percolated anchor on NP. Kiss (in press) recognises the issue, but argues that both in-situ and extraposed modification must be permitted independently, and that the problem is counter-balanced by the empirical coverage of the approach.

Contrary to Kiss (in press), I claim that the spurious ambiguity problem is much more pervasive, necessitating a refinement of Kiss's original proposal.

- (11) Man beschimpfte [**den Boten, der** den Befehl
 one insulted the messenger who the command
 überbrachte].
 delivered

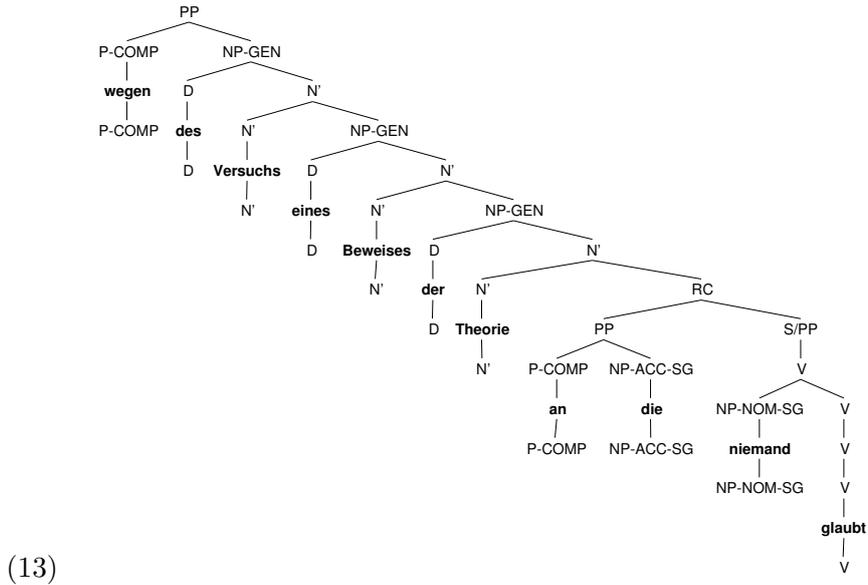
'The messenger was insulted who delivered the command.'

- (12) Man muß nicht gleich [wegen **eines**
 one must not immediately because.of an
Sachverständigen, der seinen Mund nicht halten kann],
 expert who his mouth not keep can
 die gesamte Konkurrenz kopfscheu machen.
 the whole competitors nervous make

'It does not make sense to confuse one's competitors just because an expert could not keep quiet.'

The sentences in (11) and (12), which correspond quite closely to (8) and (9), respectively, illustrate that head-initial structures are another source for spurious ambiguity.

Again, Generalised Modification permits restriction of the same index at several points along the percolation path. As illustrated by the tree representation of the complex PP *wegen des Versuchs eines Beweises der Theorie an die niemand glaubt* 'because of the attempt at a proof of the theory, which noone believes in' given in (13), the anchor of the head noun *Theorie* 'theory' can in principle be retrieved at any dominating NP, N', or PP node, leading to massive spurious ambiguity.



If we depart from the assumption that in-situ modification must always be permitted, a solution to the problem will involve further specification of the conditions under which a percolated anchor can be accessed.

What we need to do is, first, draw a distinction between a local anchor and percolated anchors, and second, distinguish inherited anchors according to whether they are already available for modification or not. Availability of inherited anchors will then depend on the intervention of other syntactic material between the modifier and the semantic target of modification. To implement this, it will be necessary to divide ANCHORS into two lists⁵, ACTIVE and INERT.

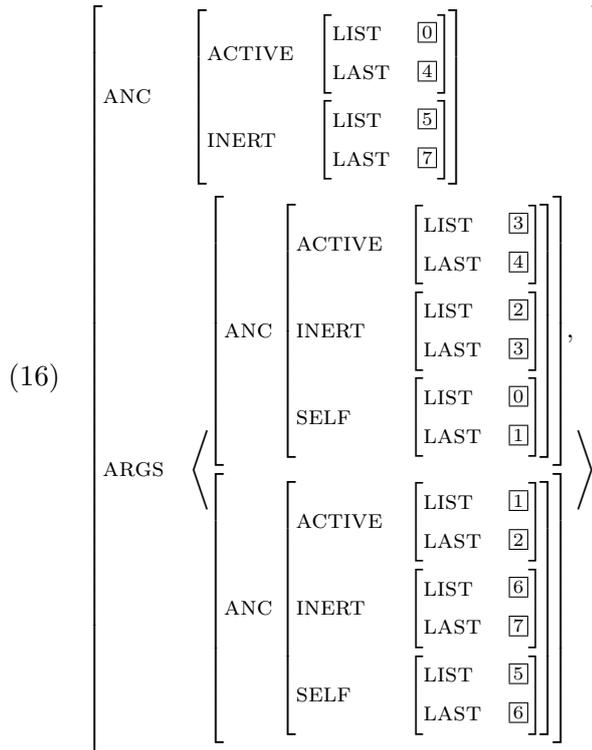
$$(14) \left[\text{SYNSEM} \mid \text{LOC} \mid \text{ANC} \begin{bmatrix} \text{ACTIVE} & *diff\text{-list}* \\ \text{INERT} & *diff\text{-list}* \end{bmatrix} \right]$$

Given binary branching structures, spurious ambiguity of the kind described above can only arise in two situations: either with percolation along a head projection path or with percolation at the right edge.

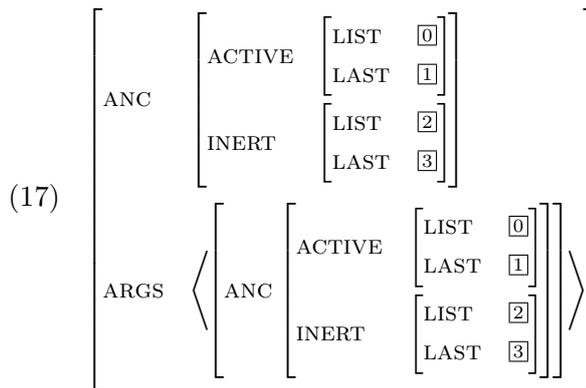
$$(15) \left[\begin{array}{l} \text{ANC} \\ \text{ARGS} \end{array} \left[\begin{array}{l} \text{ACTIVE} \left[\begin{array}{l} \text{LIST } \boxed{1} \\ \text{LAST } \boxed{4} \end{array} \right] \\ \text{INERT } \boxed{0} \end{array} \right] \left[\begin{array}{l} \text{ANC} \\ \text{ANC} \end{array} \left[\begin{array}{l} \text{ACTIVE} \left[\begin{array}{l} \text{LIST } \boxed{3} \\ \text{LAST } \boxed{4} \end{array} \right] \\ \text{INERT} \left[\begin{array}{l} \text{LIST } \boxed{2} \\ \text{LAST } \boxed{3} \end{array} \right] \end{array} \right] \left[\begin{array}{l} \text{ACTIVE} \left[\begin{array}{l} \text{LIST } \boxed{1} \\ \text{LAST } \boxed{2} \end{array} \right] \\ \text{INERT } \boxed{0} \end{array} \right] \end{array} \right] \right]$$

In order to control for the latter, all we need to do is to make composition of the ANCHORS lists sensitive to the position of the daughter from which they inherit: thus, the INERT value of the mother will correspond to the INERT value of the right daughter, whereas the ACTIVE value of the mother will be the concatenation of the ACTIVE values of the daughters with the INERT value of the left daughter. In other words, to make a percolated anchor available for modification by an extraposed phrase, it needs to be inherited, once, from a left daughter.

So far, I have only addressed the issue of anchor percolation, but we still need to specify how anchors are inserted in the first place. In order to avoid spurious ambiguity between in-situ modification and “extraposed” modification of the same head, I shall suggest that percolated anchors are never introduced below the maximal sign within a head projection. As our main concern here is relative clause extraposition, we shall concentrate on nominal anchors only. To achieve this, I shall use an additional feature SELF, again a difference list, whose single member is the anchor of the noun phrase (KEY|INDEX and KEY|HANDEL) for saturated nominal signs and the empty difference list for all other signs.⁶ We can then adapt our Anchor Percolation Principle as follows:



I have so far focussed on binary branching structures only, however, the DFKI German HPSG grammar also makes use of unary rules in the syntax, including traceless introduction of SLASH. Following the rationale adopted for binary phrase structures that only linear intervention of a (right) daughter can make an anchor active for relative clause attachment, it is clear that unary rules, in general, should not be able to activate any inert anchors, for principled reasons. The only exception amongst unary rules will be the aforementioned slash introduction rules, since extracted material will always be realised further to the left, thereby precluding spurious attachment ambiguities. Thus, we will have to distinguish slash introduction rules from ordinary unary projections. I shall propose that, in these latter structures, i.e., all unary rules other than slash introduction rules, the ANC lists are projected unaltered from daughter to mother, as captured by the following constraint:



Our specification of slash introduction rules, however, will need to take into account that, in German, extracted material realised in the *Vorfeld* can contain the antecedent of a right-extraposited relative clause in the *Nachfeld*.

- (18) [die Eltern derjenigen **Schüler**] wurden zum Direktor
 the parent of.those pupils were to.the headmaster
 zitiert, **die** auf dem Schulweg Löcher in Autoreifen
 called who on the way to school holes into car tyres
 gestochen hatten.
 pierced had

‘The parents of those pupils who had punctured car tyres on their way to school were called before the headmaster.’

- (19) [Ein **Märchen** erzählen] wollte er den Kindern, in **dem**
 a fairy tale tell wanted he the children in which
 es von übelsten Monstern und fürchterlichsten Greueln
 it of most evil monsters and most terrible atrocities
 nur so wimmelte.
 only so crawled

‘He wanted to tell the children a fairy tale that was packed with monsters of the most evil sort and with the most terrible atrocities.’

Thus, introduction of a slash dependency in the *Mittelfeld* will add the anchors contributed by an extracted phrase to the local list of active anchors:

$$(21) \quad \left[\begin{array}{l} \text{SS} \\ \text{NONLOC} \\ \text{ARGS} \end{array} \left[\begin{array}{l} \text{LOC} \left[\begin{array}{l} \text{CAT} \mid \text{HD} \\ \text{MOD} \left[\begin{array}{l} \text{rel-head} \\ \text{LOC} \left[\begin{array}{l} \text{CAT} \\ \text{CONT} \left[\begin{array}{l} \text{HD} \quad \textit{noun} \\ \text{SC} \quad \langle [] \rangle \\ \text{INDEX} \quad \boxed{z} \end{array} \right] \right] \right] \end{array} \right] \end{array} \right] \\ \text{CONT} \mid \text{KEY} \quad \boxed{k} \end{array} \right] \\ \text{NONLOC} \left[\begin{array}{l} \text{SLASH} \quad \langle ! ! \rangle \\ \text{REL} \quad \langle ! ! \rangle \end{array} \right] \\ \text{SS} \left[\begin{array}{l} \text{LOC} \quad \boxed{l} \\ \text{NONLOC} \mid \text{REL} \quad \langle ! \boxed{z} ! \rangle \end{array} \right], \\ \text{ARGS} \left\langle \begin{array}{l} \text{SS} \left[\begin{array}{l} \text{LOC} \left[\begin{array}{l} \text{CAT} \mid \text{HD} \\ \text{SC} \mid \text{KEY} \quad \boxed{k} \end{array} \right] \left[\begin{array}{l} \text{fin-verb} \\ \text{INI} \quad - \end{array} \right] \right] \\ \text{NONLOC} \mid \text{SLASH} \quad \langle ! \boxed{l} ! \rangle \end{array} \right] \right\rangle \end{array} \right] \end{array} \right]$$

While local attachment of relative clauses to an N' constituent will be taken care of by a standard Head-Adjunct schema, I shall postulate a distinct phrase structure schema for nonlocal, extraposed modifiers: as illustrated in (22), this schema largely ignores the categorial restrictions specified in the relative clause's MOD-value, thereby leaving syntactic attachment highly unrestricted. Semantic attachment simply amounts to linking the local top handle of the relative clause and the nominal index of the relative pronoun to the index and handle in the ANC|TO-BIND value of the relative clause's syntactic sister. Actual binding to a percolated anchor on ANC|ACTIVE will be performed by the retrieval rules described below.

$$(22) \left[\begin{array}{l} \text{SS | LOC} \left[\begin{array}{l} \text{ANC} \left[\begin{array}{l} \text{SELF} \langle \{ ! \} \rangle \\ \text{ACTIVE} \boxed{a} \\ \text{INERT} \boxed{i} \\ \text{TO-BIND} \left\langle \left[\begin{array}{l} \text{INDEX} \boxed{i} \\ \text{LBL} \boxed{l} \end{array} \right] \right\rangle \end{array} \right. \\ \text{HD-DTR} \boxed{h} \\ \left. \begin{array}{l} \boxed{h} \text{SS | LOC | CAT} \left[\begin{array}{l} \text{SC} \langle \rangle \\ \text{ANC} \left[\begin{array}{l} \text{ACTIVE} \boxed{a} \\ \text{INERT} \boxed{i} \\ \text{TO-BIND} \langle \rangle \end{array} \right] \end{array} \right. \\ \text{ARGS} \left\langle \left[\begin{array}{l} \text{SS | LOC} \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \text{HD} \left[\begin{array}{l} \textit{rel-head} \\ \text{MOD} \left[\text{LOC | CONT | INDEX} \boxed{i} \end{array} \right] \end{array} \right. \\ \text{SC} \langle \rangle \\ \text{CONT | LTOP} \boxed{l} \end{array} \right] \end{array} \right. \end{array} \right. \end{array} \right. \end{array} \right. \end{array} \right. \end{array} \right. \end{array} \right.$$

Given that retrieval of some member from a list can only be done in a step-by-step fashion in the underlying formalism (LKB/PET), I shall assume that attachment of an extraposed relative clause will introduce an anchor into the ANC|TO-BIND list, which will subsequently be bound to a member in INH|ANC|ACTIVE by means of recursive application of unary retrieval rules.⁷

$$(23) \left[\begin{array}{l} \text{ANC} \left[\text{TO-BIND} \langle \rangle \right] \\ \text{ARGS} \left\langle \left[\begin{array}{l} \text{ANC} \left[\begin{array}{l} \text{TO-BIND} \langle \boxed{l} \rangle \\ \text{ACTIVE} \left[\begin{array}{l} \text{LIST} \langle \boxed{l}, \dots \rangle \\ \text{LAST} \langle \rangle \end{array} \right] \end{array} \right] \end{array} \right. \end{array} \right. \end{array} \right. \end{array} \right.$$

$$(24) \left[\begin{array}{l} \text{ANC} \left[\begin{array}{l} \text{TO-BIND} \langle \boxed{1} \rangle \\ \text{ACTIVE} \left[\begin{array}{l} \text{LIST} \boxed{2} \\ \text{LAST} \boxed{3} \end{array} \right] \end{array} \right] \\ \text{ARGS} \left\langle \begin{array}{l} \text{ANC} \left[\begin{array}{l} \text{TO-BIND} \langle \boxed{1} \rangle \\ \text{ACTIVE} \left[\begin{array}{l} \text{LIST} \left[\begin{array}{l} \text{FIRST} \left[\begin{array}{l} \left[\right] \right] \\ \text{REST} \boxed{2} \end{array} \right] \\ \text{LAST} \boxed{3} \langle \rangle \end{array} \right] \end{array} \right] \end{array} \right\rangle \end{array} \right]$$

While the first schema performs the actual retrieval, i.e. binding of INDEX and HANDEL values, the second schema above will recurse over the list of active anchors, making available non-initial members for binding. The way the inheritance of anchors has been set up, structurally and superficially more recent anchors will be accessed first. Obligatoriness of retrieval can easily be ensured by requiring fully saturated sentential signs to have an empty ANC|TO-BIND list.

3. Efficiency

The integration of the approach sketched here into the German HPSG grammar at DFKI was performed within a relatively short amount of time: basic implementation and debugging was completed within a week during my stay at ESSLLI in summer 2003.

In order to assess the efficiency of the Kiss-style approach, a variant of this implementation has been created, within two days, that employs rightward movement. Both implementations make use of two-step anchor percolation to control for spurious ambiguities. They differ, though, as to the way percolation is launched and retrieved, as well as to the way upward-bounding is effected. Under the binding approach, every NP contributes a percolatable anchor and retrieval rules recursively peruse the ACTIVE list for a suitable anchor. As a consequence the length of anchor lists is not restricted. Thus, difference in semantic attachment is derived at the retrieval site.

Under the movement approach, however, different semantic attachments are derived by means of an optional unary rule that launches the nonlocal dependency. For reasons of efficiency, the length of the anchor lists has been restricted to be at most 1, meaning that multiple extrapositions with the same attachment site cannot be processed. Since introduction of this nonlocal dependency is optional, retrieval

must be obligatory, if we do not want to introduce additional spurious ambiguity. Thus, bounding nodes (as well as root nodes) require the anchor features of their daughters to be empty.

Both grammars have been evaluated against a baseline from which the ANC feature and all extrapolation-specific rule schemata have been removed.

All test runs have been performed using the July 2003 version of PET (Callmeier, 2000) without ambiguity packing. Quick check paths (Kiefer et al., 1999) were computed individually for all three grammars on the same, unrelated corpus. All runs were performed on a 2.2GHz Pentium 4M (1GB RAM) with the same limit on the maximal number of passive edges for each parse, namely 100,000. Test results were recorded and evaluated using [incr tsdb()] (Oepen and Flickinger, 1998).

In order to measure both construction-specific and general processing costs, I have used test suites both with and without relative clause extrapolation: while the TSNLP and Verbmobil test suites display zero or negligibly low numbers of extrapolated relatives, the number of extrapolated relative clauses in the manually constructed Babel regression test suite is comparatively high (roughly 4.3% as opposed to 2.8% in newspaper text; cf. Uszkoreit et al., 1998).

The results are summarised in tables I through IV: Table I summarises the coverage obtained by the individual grammars on the 3 corpora, and also provides some information about the relative complexity of the corpus (average number of *words* per item), lexical ambiguity (*lex items/words*), and syntactic ambiguity (*amb*). Tables II through IV provide the actual performance data, both in terms of average values per item (*abs*) and a factor relative to the baseline (*rel*).

Table I. Coverage & Ambiguity

Test suite	words lex items		Baseline		Anaphoric		Movement	
			cov	amb	cov	amb	cov	amb
TSNLP	5.30	13.69	74.0	1.45	74.0	1.45	74.0	1.45
VM CD15	5.18	14.95	84.6	5.04	84.7	5.04	84.4	4.80
Babel	6.76	19.93	78.9	3.14	83.2	3.59	83.2	3.59

As far as coverage is concerned, the two approaches to relative clause extrapolation produce essentially identical results on each of the three corpora, with respect to coverage and ambiguity rates.⁸ If we compare, however, the performance of the two strategies, we find that the performance losses associated with the movement approach are considerable,

Table II. Performance (average number of executed tasks per item)

Test suite	Baseline	Anaphoric		Movement	
		abs	rel	abs	rel
TSNLP	227	230	1.013	302	1.330
VM CD15	10437	10530	1.009	13739	1.316
Babel	2986	3366	1.127	4574	1.532

Table III. Performance (average parse time per item (in sec))

Test suite	Baseline	Anaphoric		Movement	
		abs	rel	abs	rel
TSNLP	.01	.01	1.1076	.02	1.486
VM CD15	.54	.59	1.105	.77	1.437
Babel	.14	.17	1.174	.23	1.654

increasing the number of executed tasks by a factor between 1.3 and 1.5. The anaphoric approach, however, features an increase in executed tasks of at most 12.7%.

Although the comparison of maximal processing cost already favours the anaphoric approach over the movement variant, a closer look at the results obtained on TSNLP and VM reveals that the Kiss-style approach is virtually cost-neutral, in terms of executed tasks, on corpora that feature low or zero occurrence of (extraposed) relatives. The residual 10% increase in processing times can safely be associated with the introduction of the ANC feature, which makes copying during uni-

Table IV. Performance (average space per item (in kB))

Test suite	Baseline	Anaphoric		Movement	
		abs	rel	abs	rel
TSNLP	273	322	1.179	439	1.608
VM CD15	15373	18980	1.235	24685	1.606
Babel	4292	5572	1.298	8128	1.894

fication more expensive, an explanation which is suggested by the corresponding increase in space consumption, as detailed in Table IV.⁹

The movement approach, by contrast, is always costly, increasing the number of executed tasks by at least one third. This result is hardly surprising, given that the anaphoric approach associates most processing costs with the retrieval site, whereas the movement approach creates local ambiguity at the point where the nonlocal dependency is launched: since retrieval rules depend on the presence of a relative clause, additional cost can only arise, under the Kiss-style approach, on test items actually containing a relative clause.

4. Beyond German

Now that we have set up a basic system that permits integration of a Kiss-style approach to relative clause extraposition while avoiding the introduction of spurious ambiguity associated with the original proposal, we may now go on and explore how the current two-step approach can be incorporated into the grammar of languages other than German. Despite the fact that relative clause extraposition is cross-linguistically a quite pervasive phenomenon, none of the large scale LKB/PET grammar implementation efforts currently provide an analysis: apparently, neither the Lingo English Resource Grammar (ERG) (Copestake and Flickinger, 2000)¹⁰, nor the German (Müller and Kasper, 2000; Crysmann, 2003) grammar developed at DFKI.¹¹ As a consequence, this phenomenon is not addressed in the HPSG Grammar Matrix (Bender et al., 2002) either, the current version of which has been distilled mainly from the Lingo ERG.

Although extraposition in German is less restricted than, e.g., in English where extraposition from topicalised constituents appears to be ruled out, English nevertheless licenses extraposition from a variety of structural contexts, e.g. subjects, objects, or even fronted wh-constituents.

- (25) a. She met a soldier at the party that she really likes.
 b. Who do you know that you can really trust? (Kiss, 2003)

Leaving aside for the moment the issue of how the introduction of anchors can or should be restricted here, it is evident that the percolation principles presented above are actually stated independently of any concrete SELF value. Thus, provided the only difference between the scope of extraposition in English and German should lie with the potential sites of anchor *introduction*, the present approach to anchor *percolation* could probably be carried over unmodified.

One of the major differences between German and English phrase structure certainly is the degree of canonicity with respect to head-dependent order. In German, we can observe a good deal of alternation here (think of pre- and postpositions), whereas English tends to be more strongly head-initial, although head-final structures can be found, most notably with Head-Subject structures. Thus, English, just like German, permits new anchors to be introduced on either daughter of a binary schema, left or right. As a consequence, the situations to be controlled for are structurally fully identical.

In our analysis of the source of the spurious ambiguity problem in German, we observed that the problem will always arise whenever local attachment and attachment to a percolated anchor may derive identical surface strings. The solution I have proposed therefore builds on a notion of linear intervention: anchors that are right-peripheral stay inert, anchors which are not, become active. The way this is stated is actually pretty much independent of any functional distinctions such as the position of the head or the type of dependent. As a result, we can directly carry over the solution developed for German to any LKB/PET grammar that only distinguishes between binary and unary rule types.¹² Having established the versatility of the current approach in a cross-lingual context, one might wonder though, whether a two-step approach to anchor percolation will always be necessary. As far as I can foresee, the division between inert and active anchors will always make sense, as long as (locally) right-peripheral nodes can introduce anchors. A language where this may not be the case will be canonically head-final, and, furthermore, will either disallow extraposition from extracted dependents, or else, not feature extraction at all. If all these criteria are met in a language, the distinction between active and inert anchors could be conflated, obviating the need for the ANC|INERT list in a grammar of that language. Of course, even under these special circumstances, the approach described here should give the right results, even if unaltered, since a proper specification of ANC|SELF values will guarantee that the ANC|INERT list of right-peripheral daughters will then simply be empty. For the case of SVO languages, such as English, however, where anchors can be introduced initially and finally in a local tree, the utility of the present approach should be evident.

Before closing our discussion, let us briefly return to the difference in restrictiveness that can be observed for relative clause extraposition in German and English.

One such difference has already been addressed above: owing to the fact that extraposed vs. local attachment to *prepositional* phrases or the NPs they contain is empirically indistinguishable, the current ap-

proach already blocks the extraposed analysis as a spurious ambiguity. As postpositions are rather uncommon in English, the unavailability of extraposed structures analogous to (9) is already accounted for by the very mechanisms that rule out spurious ambiguity in (12).

The other major difference relates to extraposition from extracted constituents: in German, extraposition is possible with fronted wh-phrases as well as topicalised constituents, whereas in English, extraposition from topicalised phrases is barred.

(26) * Micro brews, I like that are located around the bay area.

(27) Den Mann hat sie gesehen, den ich gestern getroffen
 the man has she seen who I yesterday met
 hatte.
 had

‘She saw the man who I had met yesterday.’ (Kiss, 2003)

In general, extracted constituents are not entirely unproblematic for a theory of extraposition: as the latter process is known to be upward bound (Ross, 1967), i.e., typically clause-bound, and percolation of indices may not descend the tree, modification of an antecedent outside the local clause is not trivial. There are, in principle, two ways to resolve this issue: either, anchors are inserted at the gap site, or else, some other mechanism is invoked to relocate binding of the UDC into the clausal domain and percolate the anchor from there (see below).

For Kiss (2003), the first strategy is not viable, as local/extraposed attachment is formally non-distinct there, being subsumed under the principle of Generalised Modification. Thus, in traceful formalisms/theories local attachment of a relative clause to a trace cannot be controlled for, deriving the ungrammatical (28).

(28) * [Den Mann]_i hat sie [_{t_i} den ich gestern traf] gesehen
 the man has she who I yesterday met seen
 (Kiss, 2003)

Comparing the constructions featuring extraposition from fronted constituents in English and German, Kiss (2003) observed that, in both languages, it is V2 structures that permit such extraposition, whereas English topicalisation, failing to trigger subject-verb inversion, does not appear to license extraposition from fronted constituents either. He therefore suggests that the phrase structure of wh-fronting in English is much like that of V2 structures in German, which he believes not to be instances of the Head-Filler schema. Instead, he argues, the nonlocal dependency is bound by the inverted verb in second position, which

then subcategorises, via SPR, for the wh-constituent in sentence-initial position, thereby assimilating the analysis of wh-fronting in English to that of *tough*-constructions. As an important side-effect, binding of the NONLOCAL|SLASH is moved from the head-filler structure, which is actually outside the local clause onto the head of the clause. This move makes it possible to introduce the anchor of the phrase subcategorised via SPR into the ANCHORS set of the finite verb, and, accordingly, percolate it within the domain of that head, i.e. the local clause. For topicalisation in English, Kiss (2003) maintains the classic analysis in terms of head-filler structures, by-passing the head of the clause. Consequently, the filler's anchor will not be introduced in the local clause, and therefore, extraposition from topicalised constituents in English is correctly ruled out.

Although the analysis of extraposition in English proposed by Kiss (2003) is actually quite neat, one may wonder though, whether integration of the present refined proposal regarding anchor percolation will necessitate major adaptation of any grammar that features extraposition from extracted constituents? Fortunately, given the properties of the present approach (i.e., the distinction between local and extraposed attachment) and the absence of traces from the underlying formalism (LKB/PET), we are not faced with any strong commitment as to where anchors of constituents must be introduced in order to feature in nonlocal dependencies: while an analysis along the lines of Kiss (2003) will always be possible, provided the extraposition facts in a particular language can be related to the same phrase structural properties as in English or German, the approach suggested here, however, is equally compatible with anchor introduction at the gap site, thereby ensuring a high degree of portability. Taking the case of English as an example, integration of relative clause extraposition would thus involve two things: first, unification of the above schemata for anchor percolation with all binary and unary rules that are not involved in anchor retrieval, and second, specification of the conditions on anchor introduction. If a quick and painless integration is all that is called for, it will be sufficient to segregate slash-introduction (lexical) rules according to the value of the slashed element's NONLOC|QUE value: if empty, no anchor for the slashed constituent should be introduced into ANC|SELF (topicalisation), if not, the HANDEL and INDEX features of the slashed element will get inserted as an anchor by the slash introduction rule. Thus, even under such an admittedly simple scenario, the number of slash introduction rules for a language such as English would only increase from 3 to 6. Needless to say, that for a language such as German, where all UDCs pattern alike with respect to extraposition, the single added

cost in terms of grammar code lies with the attachment and retrieval rules, i.e. 3 rules in total.

5. Conclusion

Although the present approach is admittedly slightly less general than the proposal after which it is modelled (Kiss, in press) — we need two distinct phrase structure schemata instead of one — this appears to be the price to pay, if we want to avoid proliferation of spurious ambiguity in an implemented grammar of German. On the other hand, the current approach to controlling spurious ambiguity is general enough to be applicable to almost any other language that features relative clause extrapolation, since it is formulated independently of the position of the head, which is variable in German. Furthermore, we have found that besides head-projection paths it is left-branching structures which are a notorious source for spurious ambiguity in extrapolation contexts. I therefore expect the current results to be of some use for an implementation of relative clause extrapolation in more strictly head-initial languages as well. I have shown that the current proposal can be carried over unmodified to account for similar phenomena in English, without incurring any substantial changes to existing implemented grammars. Moreover, I have shown that our implementation of an anaphoric approach to relative clause extrapolation is much more efficient than the alternative account in terms of rightward movement. Taken together, the independence of head position and the ease of integration with existing grammars will hopefully make this approach qualify as a candidate for integration into language independent core grammars, such as the HPSG Matrix Grammar initiative (Bender et al., 2002).

Acknowledgements

A previous version of this paper has been presented at the ESSLI workshop on “Ideas and Strategies for Multilingual Grammar Engineering” in Vienna, August 2003. I would like to thank the audience, and, in particular, Dan Flickinger, Melanie Siegel, Tracy Holloway King and Stefan Müller for various comments on the version presented there. A great many thanks go to the two anonymous reviewers: their comments have been invaluable for strengthening and clarifying the ideas proposed in this paper. The work reported on here 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.

Notes

¹ In their corpus study of the 10,000 structurally annotated sentences from the NEGRA 1.0 corpus, Uszkoreit et al. (1998) find that 24% of all relative clauses are actually extraposed, meaning that 2.8% of all sentences contain an extraposed relative clause. Furthermore, they show that extraposition is almost obligatory (>90%), if only one lexical item intervenes between the relative clause and its antecedent.

² An exception here is Stefan Müller's Babel system which permits direct processing of discontinuous constituents and for which also a large-scale grammar is available (Müller, 1996; Müller, 2004). With respect to performance, however, the Babel system cannot currently compete with systems such as PET in the context of real-world applications: Comparing the relative performance differences between PET and Babel, (Müller, 2004) reports that his PET grammar performed around 13 times faster than the Babel grammar, when parsing the Verbmobil corpus, although these figures should be handled with caution, given that the Babel grammar implements several constructions not covered by the LKB/Pet grammar (see, again Müller, 2004 for details). Taking into account recent performance improvements of around factor 5.5 (14.5) for the DFKI German HPSG grammar (Crysmann, 2003), we can expect an even higher factor.

³ See also Wittenburg (1987) for a similar approach to relative clause extraposition in English.

⁴ Müller (2004) cites examples of complement extraposition from complex NPs. He does not, however, offer any explanation for the clear-cut difference in acceptability observable with the data in (1a) and (2a). Thus, even though acceptable examples of complement extraposition may exist, the contrasts between (1) and (2) clearly show that the phenomena are actually quite distinct. Note further, that the examples in (2) become perfectly acceptable, if the complement clause is realised in situ. Acceptable examples of complement extraposition from adjuncts have so far not been reported in the literature.

⁵ Owing to the absence of sets in LKB and PET, I use difference lists to percolate bags of anchors. The issue of anchor retrieval, i.e., the binding of relative clauses to different anchors on this list, will be addressed at the end of this section.

⁶ Technically, this distinction is enforced on the level of phrase structure rule types (NP rules vs. all others), together with a constraint on lexical signs to have an empty SELF list.

⁷ It is hoped that retrieval rules will at some point become obsolete once the MRS implementation in LKB supports underspecification for intersective modifier attachment. Such underspecification is independently required to ensure a sound and efficient treatment of intersective modifiers in German (Crysmann, to appear).

⁸ The losses in coverage on VM-CD15 observable with the movement approach are essentially due to an exhaustion of available resources, viz. the limit of 100,000 passive edges.

⁹ In order to substantiate this claim and to measure the impact the additional list-valued feature have on parsing efficiency, I have performed a comparison between the baseline and a special version of the anaphoric grammar, where the rule schemata for

relative clause extraposition have been disabled, but feature percolation is performed as usual. On VM data, this grammar uses 0.9% more tasks, 10.5% more time and 23.5% more space than the baseline. Closer inspection of the performance data revealed that the number of copies performed per second was roughly 10% lower than the baseline.

¹⁰ As pointed out to me by Dan Flickinger (p.c.), a treatment of extraposition is currently being developed for the Lingo ERG, although this implementation effort has so far not been merged with the main development branch.

¹¹ The only implementation done so far in the context of large scale HPSG grammar engineering is Stefan Müller's Babel grammar (Müller, 2004). However, the implementation there builds on the availability of discontinuity, meaning that the analysis cannot be carried over to formalisms which assume continuous constituents only.

¹² The extension to ternary rule types should be trivial, though, as every non-right daughter will actually pattern with the left daughter in a binary rule.

References

- Bender, E. M., D. Flickinger, and S. Oepen: 2002, 'The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammar'. In: J. Carroll, N. Oostdijk, and R. Sutcliffe (eds.): *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*. pp. 8–14.
- Callmeier, U.: 2000, 'PET — A Platform for Experimentation with Efficient HPSG Processing Techniques'. *Journal of Natural Language Engineering* **6**(1), 99–108.
- Copestake, A.: 2001, *Implementing Typed Feature Structure Grammars*. Stanford: CSLI Publications.
- Copestake, A. and D. Flickinger: 2000, 'An open-source grammar development environment and broad-coverage English grammar using HPSG'. In: *Proceedings of the Second conference on Language Resources and Evaluation (LREC-2000)*. Athens.
- Crysmann, B.: 2003, 'On the Efficient Implementation of German Verb Placement in HPSG'. In: *Proceedings of RANLP 2003*. Borovets, Bulgaria, pp. 112–116.
- Crysmann, B.: to appear, 'Underspecification of Intersective Modifier Attachment: Some Arguments from German'. In: S. Müller (ed.): *Proceedings of the 11th Int'l Conference on HPSG, Aug 3–6, 2004, Leuven*. Stanford, CSLI.
- Gibson, E.: 1998, 'Linguistic Complexity: Locality of Syntactic Dependencies'. *Cognition* **68**(1), 1–76.
- Hawkins, J. A.: 1994, *A Performance Theory of Order and Constituency*. Cambridge University Press.
- Kathol, A.: 1995, 'Linearization-Based German Syntax'. Ph.D. thesis, Ohio State University.
- Kathol, A.: 2000, *Linear Syntax*. Oxford: Oxford University Press.
- Kathol, A. and C. Pollard: 1995, 'Extraposition via Complex Domain Formation'. ms., Ohio State University.
- Keller, F.: 1995, 'Towards an Account of Extraposition in HPSG'. In: *Proceedings of the Ninth Meeting of the European ACL*. Dublin.
- Kiefer, B., H.-U. Krieger, J. Carroll, and R. Malouf: 1999, 'A Bag of Useful Techniques for Efficient and Robust Parsing'. In: *Proceedings of the 37th An-*

- nual Meeting of the Association for Computational Linguistics (ACL-99)*. pp. 473–480.
- Kiss, K.: 1981, ‘Structural Relations in Hungarian, a “free” word order language’. *Linguistic Inquiry* **12**, 185–213.
- Kiss, T.: 2003, ‘Phrasal Typology and the Interaction of Topicalization, Wh-movement, and Extraposition’. In: J.-B. Kim and S. 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, T.: in press, ‘Semantic Constraints on Relative Clause Extraposition’. *Natural Language and Linguistic Theory* **forthcoming**.
- Kuno, S.: 1974, ‘The Position of Relative Clauses and Conjunctions’. *Linguistic Inquiry* **5**, 117–136.
- Lewis, R.: 1996, ‘A Theory of Grammatical but Unacceptable Embeddings’. *Journal of Psycholinguistic Research* **25**, 93–116.
- MacWhinney, B. and C. Pleh: 1988, ‘The Processing of Restricted Relative Clauses in Hungarian’. *Cognition* **29**, 95–141.
- Müller, S.: 1996, ‘The Babel-System: An HPSG Prolog Implementation’. In: *Proceedings of the 4th International Conference on the Practical Application of Prolog*. London, pp. 263–277.
- Müller, S.: 1999, *Deutsche Syntax — deklarativ*, Linguistische Arbeiten. Tübingen: Niemeyer.
- Müller, S.: 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, S. and W. Kasper: 2000, ‘HPSG Analysis of German’. In: W. Wahlster (ed.): *VerbMobil: Foundations of Speech-to-Speech Translation*. Berlin: Springer, pp. 238–253.
- Nerbonne, J.: 1994, ‘Partial Verb Phrases and Spurious Ambiguities’. In: J. Nerbonne, K. Netter, and C. Pollard (eds.): *German in Head-Driven Phrase Structure Grammar*, No. 46 in Lecture Notes. Stanford University: CSLI Publications, pp. 109–150.
- Oepen, S. and D. Flickinger: 1998, ‘Towards Systematic Grammar Profiling. Test Suite Technology Ten Years After’. *Journal of Computer Speech and Language* **12**, 411–436.
- Ross, J. R.: 1967, ‘Constraints on Variables in Syntax’. Ph.D. thesis, MIT.
- Uszkoreit, H., T. Brants, D. Duchier, B. Krenn, L. Konieczny, S. Oepen, and W. Skut: 1998, ‘Studien zur performanzorientierten Linguistik. Aspekte der Relativsatzextraposition im Deutschen’. *Kognitionswissenschaft* **7**, 129–133.
- Wittenburg, K.: 1987, ‘Extraposition from NP as Anaphora’. In: G. Huck and A. Ojeda (eds.): *Discontinuous Constituency*, No. 20 in Syntax and Semantics. New York: Academic Press, pp. 428–445.