Yet Another Paper about Partial Verb Phrase Fronting in German

Stefan Müller*
Humboldt University Berlin
Philosophische Fakultät II
Institut für deutsche Sprache und Linguistik
Lehrstuhl Computerlinguistik
Jägerstraße 10/11
D-10099 Berlin
stefan@compling.hu-berlin.de

Abstract

I describe a very simple HPSG analysis for partial verb phrase fronting. I will argue that the presented account is more adequate than others made during the past years because it allows the description of constituents in fronted positions with their modifier remaining in the non-fronted part of the sentence.

A problem with ill-formed signs that are admitted by all HPSG accounts for partial verb phrase fronting known so far will be explained and a solution will be suggested that uses the difference between combinatoric relations of signs and their representation in word order domains.

1 Introduction

During the last years, several different analyses for partial verb phrase fronting have been proposed (Pollard, To appear; Nerbonne, 1994; Baker, 1994; Hinrichs and Nakazawa, 1994b). The most promising account so far has been the one of Hinrichs and Nakazawa. This account, however, suffers from some drawbacks that will be discussed in section 4. I will present a rather simple account that uses the standard NONLOC mechanism HPSG (Pollard and Sag, 1994) provides. In section 3.3, I will discuss a problem that arises for all accounts of partial verb phrase fronting: underspecified COMPS lists. By the means of a new daughter (licensing daughter) in a schema for the introduction of nonlocal dependencies this problem will be solved.

*This paper is available via the WWW: http://www.compling.hu-berlin.de/~stefan/Pub/e_pvp.html Thanks to Frank Keller for comments on earlier versions of this paper.

2 The Phenomena

In German, it is possible to front non-maximal verbal projections.[1]

(1) a. [Erzählen] wird er seiner Tochter
tell will he his daughter
ein Märchen.
a fairy tail
'He will tell his daughter a fairy tale.'

b. [Erzählen müssen] wird er
tell must will he
seiner Tochter ein Märchen.
his daughter a fairy tale
'He will have to tell his daughter a fairy tale.'

In a series of papers, Hinrichs and Nakazawa argued for a special rule schema that combines the verbs of a so-called *verbal complex* before the arguments of the involved verbs are combined with the verbal complex. Because the verbal complex is build before any nonverbal argument of a verb gets saturated, it is possible to account for phenomena like *auxiliary flip*. As the verbal complex is analyzed as a constituent, the fronting of *erzählen müssen* in (1b) can be explained as well. There is no problem with sentences like those in (1) for the standard NONLOC mechanism. *Erzählen müssen* is a constituent in the non-fronted position in (2) and the same holds if the verbal complex is fronted.

(2) Er wird seiner Tochter ein Märchen [erzählen müssen].

There are, however, examples where a partly saturated verbal complex is fronted.

(3) a. [Seiner Tochter ein Märchen erzählen] wird er.

b. [Ein Märchen erzählen] wird er seiner Tochter.

c. [Ein Märchen erzählen] wird er seiner Tochter müssen.

[1]The examples (1) and (3) are taken from Hinrichs and Nakazawa (1994b).

d. [Seiner Tochter erzählen] wird er das Märchen.

A verb with some of its arguments may appear in the *Vorfeld* leaving other arguments in the *Mittelfeld*.

As (4) shows, it is possible that a PP in the *Mittelfeld* modifies a fronted verbal complex.

(4) Den Kanzlerkandidaten ermorden
the chancellor.candidate kill
wollte die Frau mit diesem Messer.
wanted the woman with this knife
'The woman wanted to kill the candidate with this knife.'

Sentences like (5a) are ungrammatical. It is not possible to front parts of the verbal complex that would be located in the middle of the verbal complex in a verb final sentence (5b).

(5) a. * Müssen wird er ihr
must will he her
ein Märchen erzählen.
a story tell

b. , weil er ihr ein Märchen erzählen müssen wird.

3 The Analysis

3.1 Basic Assumptions

In what follows, I assume a version of HPSG that deviates from standard HPSG in that the surface string of a phrasal sign is not determined by a relation that relates the PHON values of a sign to the PHON values of its daughters (Pollard and Sag, 1987, p. 169). Instead I will follow Reape's (1994) approach. Reape assumes word order domains as an additional level of representation. In such a domain, all daughters of a head occur. These domains differ from the daughter list in that the elements in a domain (signs) correspond in their serialization to the surface order of the words in the string. LP-constraints apply to elements of the order domain. Another basic assumption of Reape is that constituents may be discontinuous.

As Hinrichs and Nakazawa (1994a) have shown, it is reasonable to assume in addition to the head complement schema a schema that licenses the verbal complex. Hinrichs and Nakazawa introduced the concept of argument attraction into the HPSG framework. If a verbal complex is build two verbs are combined and the resulting sign inherits all arguments from both verbs. In their paper, Hinrichs and Nakazawa treat verbal complements as ordinary complements that are included in the COMPS list of their heads. It has however proven to be useful to distinguish the verbal complement from other complements (Rentier, 1994a; Müller, 1995a). The merits of this move will be discussed shortly. For the purpose of representing the information about verbal complements, the feature VCOMP is introduced. Its value is a *synsem*-object if the verb embeds another verb and *none* otherwise. The entry in the stem lexicon for the future tense auxiliary *werden* (*will*) is shown in (6). From this stem the morphology component produces the finite form shown in (7). In German, almost any complement of a verb can be fronted, subjects as well as objects. Therefore, for finite forms the subject is included into the COMPS list, from where extraction is possible. For nonfinite forms the subject does not appear on COMPS but stays in the SUBJ list.[2] Schema 1 licenses verb cluster structures.[3] A head is combined with its verbal complement ([1]). The resulting sign is a

werden:

$$\left[\begin{array}{l} \text{HEAD} \quad \left[\begin{array}{l}\text{SUBJ}\ \boxed{1}\\ \textit{verb}\end{array}\right] \\ \text{COMPS}\ \boxed{2} \\ \text{VCOMP}\ \text{V}[\text{LEX}+,\textit{bse},\text{SUBJ}\ \boxed{1},\text{COMPS}\ \boxed{2},\\ \qquad\qquad \text{VCOMP}\ \textit{none}] \\ \textit{cat}\end{array}\right] \tag{6}$$

wird:

$$\left[\begin{array}{l} \text{HEAD} \quad \left[\begin{array}{l}\text{VFORM}\ \textit{fin}\\ \text{SUBJ}\ \langle\rangle\\ \textit{verb}\end{array}\right] \\ \text{COMPS}\ \boxed{1} \oplus \boxed{2} \\ \text{VCOMP}\ \text{V}[\text{LEX}+,\textit{bse},\text{SUBJ}\ \boxed{1},\ \text{COMPS}\ \boxed{2},\\ \qquad\qquad \text{VCOMP}\ \textit{none}] \\ \textit{cat}\end{array}\right] \tag{7}$$

Schema 1 (Verb Cluster Schema)

$$\left[\begin{array}{ll} \text{SYNSEM} & \left[\begin{array}{ll}\text{LOC}|\text{CAT} & \left[\text{VCOMP}\ \textit{none}\right]\\ \text{LEX} & +\end{array}\right] \\ \text{DTRS} & \left[\begin{array}{ll}\text{HEAD-DTR} & \left[\begin{array}{ll}\ldots|\text{VCOMP} & \boxed{1}\\ \text{DOM} & \boxed{2}\end{array}\right]\\ \text{CLUSTER-DTR} & \left[\begin{array}{ll}\text{SYNSEM} & \boxed{1}\\ \text{DOM} & \boxed{3}\end{array}\right]\\ \text{COMP-DTRS} & \langle\rangle\\ & \textit{head-cluster-structure}\end{array}\right] \\ \text{DOM} & \boxed{2} \bigcirc \boxed{3} \\ & \textit{phrasal-sign}\end{array}\right]$$

[2] see (Kiss, 1993) for details

[3] I will not go into the details of the domain formation in verb cluster structures. For details see (Müller, 1995b).

verbal complex or a part of a verbal complex. It is marked LEX+ because it can in turn be embedded.

(8) , weil er ihm ein Märchen
because he him a fairy tale
[[erzählen lassen] hat].
tell let has
'because he has let him tell the story.'

3.2 The LEX Feature

The LEX feature in the entry for *werden* ensures that a matrix verb is combined with its verbal complement before the verbal complement is saturated by one of its complements. It is therefore possible to avoid multiple structures in the *Mittelfeld*.

(9) a. Er wird seiner Tochter ein Märchen [erzählen müssen].

b. Er wird seiner Tochter [[ein Märchen erzählen] müssen]].

c. Er wird [[seiner Tochter ein Märchen erzählen] müssen]].

But exactly those constituents that have to be avoided in the *Mittelfeld* are needed in the *Vorfeld*. Very complicate mechanisms have been introduced to cope with this problem without a lot of spurious ambiguities (Nerbonne, 1994; Hinrichs and Nakazawa, 1994b). I will suggest a solution to the problem that is very simple: If it is the case that an embedded verb or verbal complex has to be LEX+ when verb and complement are combined locally and if it is the case that this does not hold if a nonlocal dependency is involved than the simplest solution is to view LEX not as a local feature. If one assumes that LEX lives under the path SYNSEM instead of SYNSEM|LOC than the problem turns into a non-issue.

Figures 1 and 2 show the analyses of the sentences in (10).[4] In the analyses of (10a), a trace functions as a verbal complement. In (10b) a trace for a verb is modified by an adverb.

(10) a. Seiner Tochter erzählen wird er das Märchen.

b. Vortragen wird er es morgen.

Sentences like (5a) are ruled out because *wird* selects a complement in *bse*-form that has a VCOMP value *none*. As *erzählen* does not appear in any COMPS list it is not possible for the verb to count as an argument of the fronted verbal complex that is saturated in the *Mittelfeld*. This is the case in Pollards account. Hinrichs and Nakazawa have to block this case by stating type constraints on lists of attracted arguments. With a separate VCOMP feature this problem disappears.

3.3 The Problem of Underspecified COMPS Lists

In this section, I will address a problem that seems to have gone unnoticed until now. All analyses that involve argument attraction admit signs with underspecified COMPS lists. So in (1), *wird* is combined with a trace or a lexical rule is applied to it. The LOC value of the verbal complement is put into SLASH and the arguments of the verbal complement are attracted by the matrix verb. This list of arguments, however, is not instantiated in the resulting sign. It remains variable until the SLASH element becomes bound. Therefore, the HPSG principles admit any kind of combination of totally unrelated signs. Since the COMPS list of the head is variable, any constituent is a possible complement.[5] As an HPSG theory is assumed to be a set of constraints that describe well formed descriptions of linguistic objects, this is clearly not wanted. If a grammar contains phonologically empty elements (traces, relativizers, and the like) the set of ill-formed signs will be infinite because *wird* $_i$ could be combined with arbitrarily many empty elements.[6]

It is clear that we want the matrix verb to behave in a very well defined way. It shall attract exactly the arguments of the fronted verbal projection that were not saturated by this projection, i.e., the matrix verb shall perform the argument attraction that would take place in base position, abstracting away from the value of LEX. The desired effect can be reached if a rule schema is used for the introduction of nonlocal dependencies. To introduce a nonlocal dependency for a verbal complex, this schema requires an additional licensing condition to be met. The extracted element is licensed by an actually existing verbal projection in the string. When a hearer of a sentence hears the words that have to be combined with a trace or introduce the nonlocal dependency in another way, he or she has already heard the phrase actually located in the *Vorfeld*. Therefore, the information about the nonlocal dependency is present and can be used to license the extracted element. The COMPS list of the extracted element therefore is specified. The specified COMPS are attracted by the matrix verb and the COMPS list of the matrix verb therefore does not contain any variables and our theory does not admit signs that don't describe linguistic objects.

[4] In the original grammar, I use a binary branching schema for head-complement and verb cluster structures. Adjuncts and complements are inserted into the domain of their head so that word order facts are accounted for. Due to space limitations, the figures show a tree for a flat head-complement structure.

[5] The same problem exists for analysises that treat verb second as verb movement (Kiss and Wesche, 1991; Netter, 1992).

[6] For a bottom-up parser, this would mean nontermination.

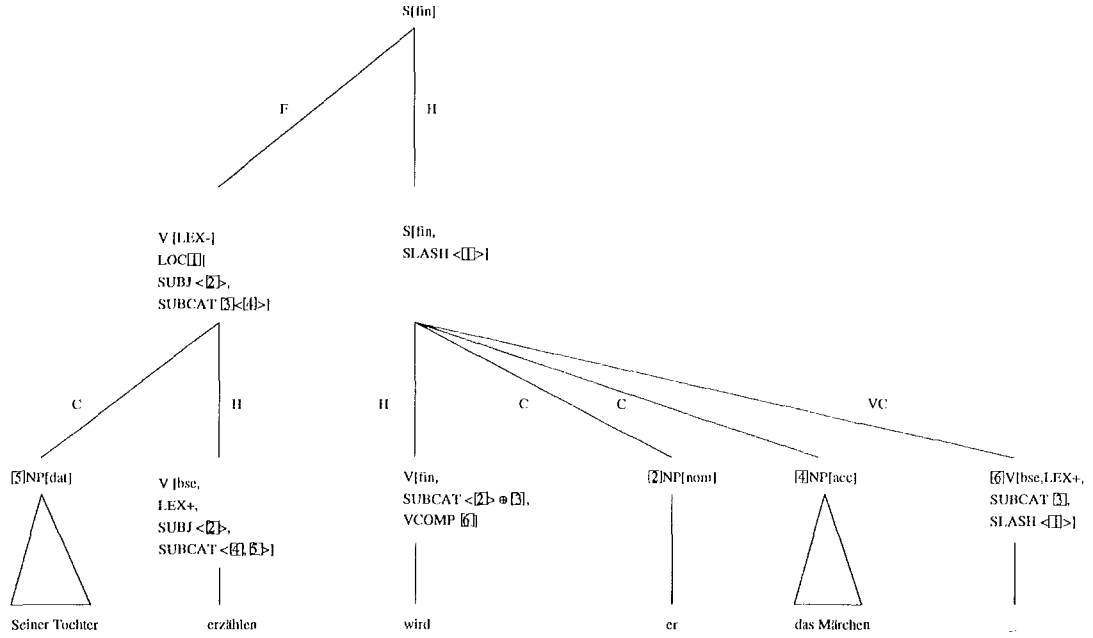


Figure 1: Analysis of *Seiner Tochter erzählen wird er das Märchen.*

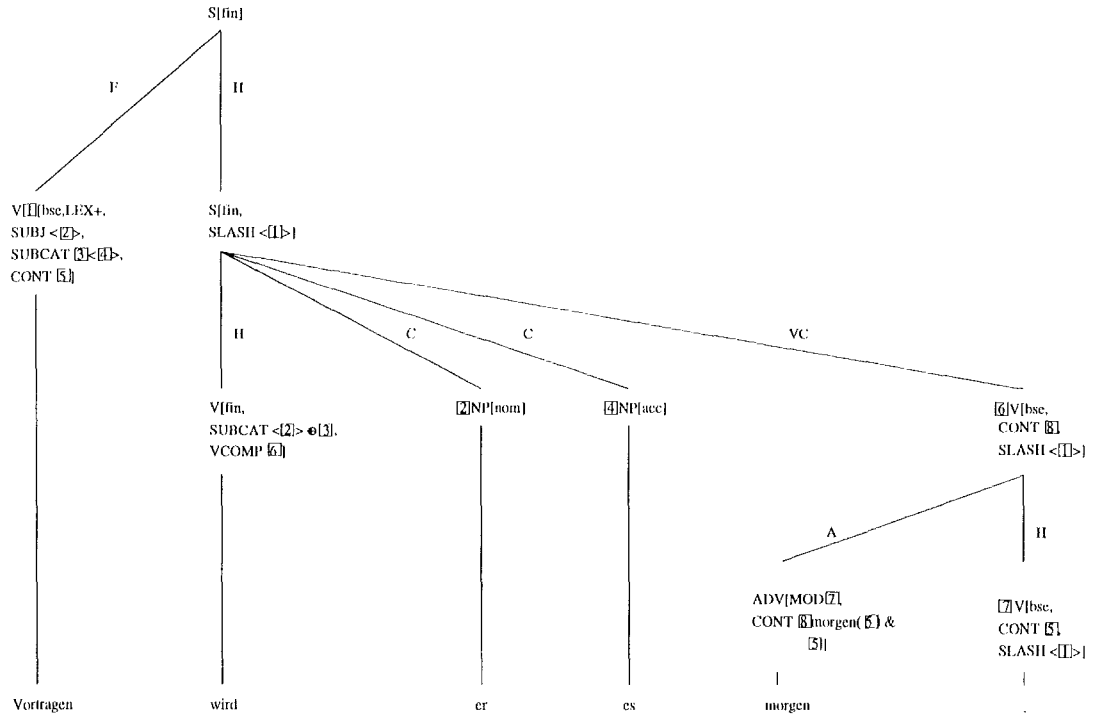
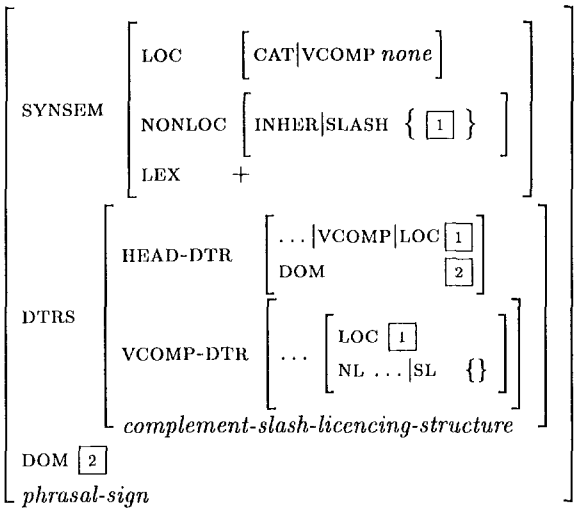


Figure 2: Analysis of *Vortragen wird er es morgen.*

Schema 2 (PVP-SLASH-Introduction-Schema)



Schema 2 shows how this is implemented. A verbal complement of a matrix verb is saturated. The VCOMP value of the resulting sign is *none*. The LOC value of the saturated verbal complement is moved into SLASH. This LOC value is licensed by another verbal projection that meets the local requirements of the matrix verb but may be positioned in the *Vorfeld*. As there are no constraints for daughters to be adjacent to each other, there may be an arbitrary number of constituents between the licensing daughter and the head daughter. The licensing daughter has licensing function only and is not inserted into the domain of the resulting sign ($\boxed{2}$) at this point of combination. However, an appropriate sign is inserted into the domain of its head when the nonlocal dependency is bound.

4 Alternatives

The drawback of the approaches of Pollard (To appear) and Nerbonne (1994) are discussed in (Hinrichs and Nakazawa, 1994b). I will not repeat the arguments against these approaches here. Instead, I will explain some of the problems of the Hinrichs and Nakazawa approach.[7]

Hinrichs and Nakazawa changed the value of SLASH into a set of signs rather than *local* objects. The fronted phrase is a maximal projection with the missing constituents moved to SLASH. The fronted partial phrase is the filler for a nonlocal dependency which is introduced by their PVP-Topicalization Lexical Rule. As SLASH elements are signs, the lexical rule can refer to the SLASH set of a SLASH element and it is thus possible to establish a relation between the COMPS list of the auxiliary and the SLASH set of the fronted verbal projection. However, the assumption that SLASH contains signs rather than local objects is a change of the basic HPSG formalism with far reaching consequences that is not really needed and that has some side effects.

In the following, I discuss two problems for this approach. Firstly, it is not possible to account for cases where a modifier in the *Mittelfeld* modifies the fronted verbal projection without assuming an infinite lexicon because the only way for a modifier to stay in the *Mittelfeld* while the modified constituent is fronted is that the modifier is contained in the SLASH set of the fronted constituent. It therefore had to be a member of the COMPS list. An infinite lexicon is both not very nice from a conceptual point of view and an implementational problem. Without a complex control strategy (late evaluation) it is not possible to implement an infinite lexicon. Another problem that was pointed out by Hinrichs and Nakazawa themselves is sentences like (11).

(11) * Gewußt, daß Peter $_i$ schlägt, habe
known that Peter hit have
ich sie$_i$.
I her
'I knew that Peter hit her.'

In (11), *sie* is extracted from the complement sentence of *gewußt* and than inserted into the COMPS list of *habe* and saturated in the *Mittelfeld*. The same problem arises for other constructions involving nonlocal dependencies.[8]

(12) a. [Da]$_i$ hatte Karl [$_i$ mit] gerechnet.
this had Karl with counted
'Karl expected this.'

b. * [[$_i$ mit] gerechnet] hatte [da]$_i$ Karl.

(13) a. Bus$_i$ will Karl [$_i$ fahren].
bus wants Karl drive
'Karl wants to go by bus.'

b. * [$_i$ fahren] will Karl bus$_i$.

5 Conclusion

A very simple solution for the PVP problem was found. A minor change in the feature geometry of signs was sufficient to cope with the spurious ambiguity problem of Pollard's (To appear) account. The account argued for in this paper can describe the fronting phenomena without the assumption of an infinite lexicon. A solution for the problem of underspecified COMPS lists was found. This solution makes use of a schema to introduce the nonlocal dependency. An introduced nonlocal dependency is licensed by an actually present element in the syntax analysis of a string. At the

[7] Due to space limitations, I cannot give a detailed discussion of their approach here. The interested reader is referred to (Müller, 1996).

[8] For an analysis of stranded prepositions in terms of nonlocal dependencies see (Rentier, 1994b) and (Müller, 1995c).

point of combination, this element plays a licensing role only and does not appear in the surface string of the build sign. This is possible because two different levels of representation for combinatorial and order information are used.

The analysis is part of an implemented fragment of German (Müller, 1996).

References

Baker. 1994. An integrated account of 'modal flip' and partial verb phrase fronting in German. appears in CLS 30-I: Papers from the Thirtieth Regional Meeting, Chicago Linguistic Society.

Hinrichs and Nakazawa. 1994a. Linearizing AUXs in German Verbal Complexes. In Nerbonne, Netter, and Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, CSLI Lecture Notes, Number 46, pages 11–38. Center for the Study of Language and Information, Stanford.

Hinrichs and Nakazawa. 1994b. Partial-VP and Split-NP Topicalization in German– An HPSG Analysis. Arbeitspapiere des Sonderforschungsbereiches 340, Bericht Nr. 58. Universität Tübingen.

Kiss and Wesche. 1991. Verb Order and Head Movement. In Herzog and Rollinger, editors, *Text Understanding in LILOG*, Lecture Notes in Artificial Intelligence, Number 546, pages 216–242. Springer-Verlag, Berlin Heidelberg New York.

Kiss. 1993. Infinite Komplementation – Neue Studien zum deutschen Verbum infinitum. Arbeiten des Sonderforschungsbereiches 282 Nummer 42, Bergische Universität Gesamthochschule Wuppertal.

Müller. 1995a. Complement Extraction Lexical Rules and Argument Attraction. Unpubl. ms., Humboldt Universität zu Berlin.

Müller. 1995b. *Head-Driven Phrase Structure Grammar für das Deutsche.* Vorlesungsskripte Computerlinguistik. Humboldt Universität zu Berlin. http://www.compling.hu-berlin.de/~stefan/Pub/e_hpsg.html.

Müller. 1995c. Scrambling in German – Extraction into the *Mittelfeld*. In *Proceedings of the tenth Pacific Asia Conference on Language, Information and Computation.* City University of Hong Kong. http://www.compling.hu-berlin.de/~stefan/Pub/e_scrambling.html.

Müller. 1996. Yet another Paper about Partial Verb Phrase Fronting in German. Humboldt-Univerität zu Berlin, http://www.compling.hu-berlin.de/~stefan/Pub/e_pvp.html.

Müller. 1996. The Babel-System–An HPSG Prolog Implementation, Overview. In *Proceedings of the Fourth International Conference on the Practical Application of Prolog,* London. http://www.compling.hu-berlin.de/~stefan/Pub/e_babel.html.

Nerbonne. 1994. Partial Verb Phrases and Spurious Ambiguities. In Nerbonne, Netter, and Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, CSLI Lecture Notes, Number 46, pages 109–150. Center for the Study of Language and Information, Stanford.

Netter. 1992. On Non-Head Non-Movement. An HPSG Treatment of Finite Verb Position in German. In Görz, editor, *Konvens 92*, pages 218–227, Berlin, Heidelberg, New York. Springer-Verlag.

Pollard and Sag. 1987. *Information-Based Syntax and Semantics.* CSLI Lecture Notes, Number 13. Center for the Study of Language and Information, Stanford.

Pollard and Sag. 1994. *Head-Driven Phrase Structure Grammar.* University of Chicago Press.

Pollard. 1994. Toward a Unified Account of Passive in German. In Nerbonne, Netter, and Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, CSLI Lecture Notes, Number 46, pages 273–296. Center for the Study of Language and Information, Stanford.

Pollard. To appear. On Head Non-Movement. In Horck and Sijtsma, editors, *Discontinuous Constituency.* Mouton de Gruyter, Berlin, New York.

Reape. 1994. Domain Union and Word Order Variation in German. In Nerbonne, Netter, and Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, CSLI Lecture Notes, Number 46, pages 151–198. Center for the Study of Language and Information, Stanford.

Rentier. 1994a. A Lexicalist Approach to Dutch Cross Serial Dependencies. In *Proceedings of Coling 94*, Kyoto.

Rentier. 1994b. A Lexicalist Approach to Dutch Preposition Stranding. In Trost, editor, *Tagungsband Konvens 94, Verarbeitung natürlicher Sprache*, pages 280–289.