Arbeitspapiere des SFB 340
Bericht Nr. 00, Xxxx 19??

LEXICALIZING THE LEFT PERIPHERY OF GERMAN FINITE SENTENCES

FRANK RICHTER, MANFRED SAILER

1. Introduction

In this paper, we discuss a new analysis of the left periphery of German finite sentences. With the term left periphery we refer to that region of a finite sentence which contains sentence initial finite verbs that might either be preceded by a constituent (verb second sentences, V2) or not (verb first sentences, V1), or sentence initial complementizers that might either be lexical or complex (verb final sentences, VF). Our analysis builds on empirical generalizations previously formulated in the topological fields model of German sentence structure. It accounts for the movement of finite verbs without lexical rules, traces nor a feature INV(erted), and eliminates the HEAD-FILLER-SCHEMA for unbounded dependencies. Its main idea is to lexicalize the left periphery or, in other words, to explain the topological makeup of the left periphery of finite sentences in terms of certain properties of a class of lexical elements. An eminent characteristic of these lexical elements will be that exactly one of them must introduce each finite sentence. The main arguments favoring our new perspective will be drawn from well known but puzzling facts about the distribution of inflectional morphems at the left periphery of finite sentences in a number of Germanic languages, and from observations that reveal clear differences between finite verbs in sentence initial and in sentence final position. We will show that a lexical theory of the elements at the left periphery of

We would like to thank Josef Bayer, Anke Feldhaus, Tilmann Höhle, Tibor Kiss, Uli Lutz, Detmar Meurers, Stefan Müller, Gerald Penn, Craig Thiersch, and the participants of the ESSLLI 98 workshop on *Current Topics in Constraint-based Theories of Germanic Syntax* for discussion and comments, and Christian Richard for help with the English language. Very special thanks to our editor, Valia Kordoni, for her support.

ISSN 0947-6954/??
© 19?? Die AutorInnen

finite sentences allows natural empirical generalizations that nonlexical or phrase structure-based accounts have not been able to provide.

The syntactic analysis will be complemented by a concrete proposal of how to provide logical forms for the syntactic structures. As semantic representation language, we employ a higher order type-theoretic language, Ty2, which is widely accepted as a standard in natural language semantics, and exploit techniques first pioneered by theories of underspecified semantics for the formulation of semantic composition rules that derive the meaning of each phrase from the meaning of its daughters.

2. Empirical Motivation

2.1. **Generalizing over Topological Fields.** Many important insights into the properties of finite sentences in German can and have been formulated in terms of the traditional, descriptively oriented topological fields model of German sentence structure. It distinguishes three types of sentences, VF, V1, and V2, and provides a schema for the classification and analysis of every finite sentence.[1] VF clauses are devided into five topological fields:

C	X	VK	Y
ob	wohl jemand den Hund	füttern würde	morgen abend?
daß	du mir ja nicht den Hund	fütterst!	
in dessen Hütte	eine Katze	gefunden wurde	
wer (daß)		kommt	

Traditionally, the X field is called Mittelfeld, VK is the verb cluster, and Y is called Nachfeld. We can distinguish three groups of VF clauses by the kinds of elements that appear in the C field, which may be lexical complementizers such as *ob* (*whether*) or *daß* (*that*) (c-VF), relative constituents such as *in dessen Hütte* (*in whose hut*) (d-VF), or w-constituents such as *wer* (*who*) (w-VF). The topology of V1 clauses differs from VF clauses only in the first field, Finit, which replaces the C field of VF clauses and contains a finite verb.

FINIT	X	VK	Y
würde	denn jemand den Hund	füttern	morgen abend
kommen	die Leute denn auch,		die du eingeladen hast
hättest	du mir doch	zugehört	damals
gib	mir mal einer den Hammer!		

In V2 clauses, one constituent precedes the Finit field. The K field is usually called Vorfeld.

[1] We adopt the formulation of the topological fields model outlined in Höhle 1986.

K	FINIT	X	VK	Y
die Frau	hat	das Buch	gelesen	gestern abend
es	hört	schon	auf	
wer	kommt?			
dann	hör	doch	zu!	

In the topological fields model of German sentence structure, all three clause types share the fields labelled X, VK and Y. They differ, however, in both the number and name of the sentence intial field(s). Our analysis will generalize over the left periphery, which will lead to a uniform account for all clause types.

Opener	Mittelfeld	Verb Cluster	Nachfeld
C	X	VK	Y
FINIT	X	VK	Y
K FINIT	X	VK	Y

In the table above, the field(s) which precede the X field in each sentence type are subsumed under a new label, *Opener*. In the following sections, we will investigate the properties of the elements in the Opener region of finite sentences and propose a unified lexicalist analysis.

2.2. **Two Arguments.** In this section, we will discuss two sets of data that are problematic to phrase structure oriented approaches in generative grammar and, as will be discussed in section 5, to previous HPSG analyses. The analysis we propose below will capture the data by a shift to a lexical treatment of the Germanic left periphery. We will argue that the data substantiate the following three claims:

1. Fronted finite verbs differ from non-fronted finite verbs (in form and meaning).
2. Fronted finite verbs pattern with complementizers (in form and meaning).
3. Fronted w-phrases in w-VF clauses pattern with complementizers.

2.2.1. *Inflection.* As mentioned in Höhle 1997, in Old English, Middle High German, Middle Low German, Modern Dutch and in some modern German dialects, the morphology of a fronted verb in V2 clauses depends on the kind of constituent that occupies the Vorfeld. As shown in (1), in Dutch V1 and V2 clauses with a second person singular subject, the verb has a special inflection if there is no Vorfeld or a non-subject stands in the Vorfeld.

(1) a. Je **hebt/*heb** het boek gekocht.
you (2sg.) have the book bought

'You have bought the book.'

b. Het boek **heb/*hebt** je gekocht.
'You have bought the book.'

c. **Heb/*Hebt** je het boek gekocht?
'Have you bought the book?'

d. Dat je het boek gekocht **hebt/*heb**.
'that you have bought the book.'

The data in (1) confirm our first claim, which predicts differences in form between fronted and non-fronted finite verbs.

In dialects with inversion forms and inflected complementizers like Overijssel, the complementizer follows the inversion paradigm. The relevant data are, again, cited from Höhle 1997.

(2) a. wy/ zy kun**t**
we/ they can-**t**

b. kunn**e** wy/ *zy
can-**e** we/ they

(3) a. azz**e** wy de trf niet verkopen kunt
that-**e** we the turf not sell can
'that we cannot sell the turf.'

b. as(*ze) zy de trf niet verkopen kunt

The data in (2) and (3) motivate our second claim, which says that complementizers and fronted finite verbs exhibit the same formal properties. To substantiate our third claim, we consider data from Bavarian as given in Bayer 1984. In Bavarian, complementizers show overt inflection in the second person.[2]

(4) ob-**st** (du) noch Minga kumst.
whether-**st** (you) to Munich come
'whether you come to Munich.'

Bayer 1984 shows that, even though the subject pronoun, *du* (*you*), is optional in (4), *-st* is clearly not a (phonologically) reduced form of the

[2]According to the data in Bayer 1984, p. 252, Lower Bavarian even shows distinct inflection paradigms in the first person plural for fronted finite verbs and complementizers on the one hand, and non-fronted finite verbs on the other. The data, however, do not fully pattern with those of the Dutch dialect in (2), because the inflection is not sensitive to the particular element in the Vorfeld.

pronoun.[3] In w-VF clauses, the same affixation appears. The inflection attaches to the fronted w-phrase. Any other placement within the w-phrase is strongly ungrammatical.

(5) Du sollst sogn, ...

a. [an wäichan Schuah]-**st** (du) wuist.
[the which shoe]-**st** (you) want
'You should say which of the shoes you want.'

b. [an(-***st**) wäichan-(***st**) Schuah]-(**st**) du wuist.

These data show that the fronted w-constituent behaves as if it were a complementizer. Under the reasonable assumption that inflectional affixes can only attach to a restricted set of elements, it follows that fronted finite verbs, lexical complementizers and fronted w-constituents in w-VF clauses are in some relevant sense of a similar syntactic category. If, in addition, we assume that, normally, only words have inflectional affixes, fronted w-constituents in w-VF clauses, even though internally potentially phrasal, must be considered words in their external behavior.[4]

In some varieties of German, there is an alternative to w-VF clauses in which the fronted w-constituent co-occurs with an overt complementizer, *daß* (wc-VF clauses). In Bavarian, wc-VF data interact with complementizer inflection. If the overt complementizer is present, the inflection is realized on *daß* rather than on the w-constituent.

(6) ... [an wäichan Schuah](***-st**) daß-**st** (du) wuist.
[the which shoe](**-st**) that-**st** (you) want

There are indications that the fronted w-constituent and the lexical complementizer form a unit. Inflected complementizers and fronted w-constituents with either inflection or a lexical complementizer cannot occur without an overt clause.[5]

---

[3]Here, Bavarian differs from colloquial standard German, which allows *-ste* endings on complementizers only if there is no overt personal pronoun:

(i) Ich weiß nich', ob-**ste** (*du) des darfst.
I know not whether-**ste** (you) that may do
'I don't known whether you are allowed to do this.'

[4]This idea will be made formally explicit in section 3.1. It is similar to the proposal of Reis 1985, which locates fronted w-constituents in w-VF clauses and lexical complementizers in c-VF clauses in the same position, despite the fact that the w-constituent may be phrasal.

[5]We are grateful to Josef Bayer (p.c.) for example (7) and for discussing examples (9) and (10) with us, and to Uli Lutz for judgments and long discussion of the data.

(7) a. Otto weiß, DAß(-**st**) du kommst, aber leider nicht
Otto knows that you come but unfortunately not
WANN.
when
b. ... * aber leider nicht WANN-**st**.
c. ... * aber leider nicht WANN daß-(**st**).

(8) Otto will noch nicht wissen, WANN(-**st**) du kommst,
Otto wants yet not know when you come
sondern OB-(*__st__) (überhaupt).
but whether-(**st**) (at all)

Interestingly, inflected complementizers and fronted w-constituents with inflection or lexical complementizer can marginally appear as first conjuncts in coordination:

(9) a. wo und wann daß du das gehört hast.
where and when that you that heard have
b. ?? wo daß und wann daß du das gehört hast.
c. * wo daß und wann du das gehört hast.

(10) a. wo und wann-**st** du das gehört hast.
where and when that you that heard have
b. ?? wo-**st** und wann**st** du das gehört hast.
c. * wo-**st** und wann du das gehört hast.

(9c) and (10c) would hypothetically have to be analyzed as coordinated sentences, where the first conjunct contains an ellipsis, as indicated in (11a). As had been shown in (7), however, complementizers of the form in (c) may not appear with an elided clause. Examples (9b) and (10b), even though not fully acceptable, are clearly better than (9c) and (10c). For these examples, an alternative analysis is needed. The option of coordinating the complementizers seems to be that alternative (11b). The availability of complementizer coordination in German is very restricted; in particular, it is impossible in (9c) and (10c).

(11) a. * $[_{\bar{S}}$wo daß $e_i]$ und $[_{\bar{S}}$ wann daß $S_i]$
b. [[wo daß] und [wann daß]] S

We conclude that (9b) and (10b), as far as they are marginally grammatical, are indeed instances of complementizer coordination. This indicates that the fronted w-constituent forms a unit with the lexical complementizer or the inflection respectively.

The data with a (possibly inflected) lexical complementizer in the first conjunct, as in (12), also point to the existence of complementizer coordination and, thus, show that fronted w-constituents form a unit with the inflection.

(12) a. *? ob und wann-**st** du das gemacht hast.
whether and when-**st** you that done have
b. ?? ob-**st** und wann-**st** du das gemacht hast.
c. * ob-**st** und wann du das gemacht hast.

The ungrammaticality of (12c) is, again, explained parallel to (8). To account for (12b), we must assume that there can be no ellipsis in the first conjunct and that the structure is licensed as complementizer coordination. But that is only possible if *wann-st* forms a (lexical) unit.[6]

2.2.2. *Verum Focus.* Besides inflection, stress facts are an empirical area that substantiates our three claims. Let us consider a particular kind of stress that results in a so-called Verum Focus reading of finite clauses (Höhle 1992). Whereas the precise semantics and pragmatics of Verum Focus remain mysterious, it seems clear that in a first intuitive characterization, Verum Focus emphasizes the truth value of the propositional content of a sentence. Empirically, Verum Focus can be distinguished from other types of focus, such as contrastive focus. What is important to our discussion is that we are able to observe the realization of one phenomenon, Verum Focus, in different clause types.

Verum Focus is achieved by stressing the finite verb in V1 and V2 clauses. In c-VF clauses, the complementizer must be stressed, and, in w-VF clauses, the w-pronoun of the w-constituent which functions as a complementizer bears main stress (Höhle 1992). In (13), stress is indicated by italics, and Verum Focus is entailed. The sentences marked with '#' lack a Verum Focus interpretation.

(13) a. Hans *hat* das Buch gekauft.
Hans has the book bought
'Hans bought the book.'
Hans hat das Buch gekauft.
b. *Hat* Hans das Buch gekauft.
c. *daß* Hans das Buch gekauft hat.
daß Hans das Buch gekauft *hat*.

[6]It remains unclear why (12a) is worse than (12b), in contrast to the seemingly corresponding case in (10), and why the option of complementizer coordination is so limited and leads to relatively bad sentences (9b–12b).

d. *wess*en Buch Hans gekauft hat.
wessen Buch Hans gekauft *hat*.

The data in (13) confirm our three claims. First, Verum Focus can only be achieved by stressing a fronted verb, but not by stressing a non-fronted verb, which points to a syntactic difference between the two (claim 1). Second, lexical complementizers may receive stress to express Verum Focus just like fronted finite verbs (claim 2). Third, the w-pronoun of fronted w-constituents in w-VF clauses is stressed to achieve Verum Focus. The w-constituent patterns, thus, with lexical complementizers (claim 3).

Finally, we want to consider those German dialects in which the left periphery of w-VF clauses may consist of a w-constituent and *daß* (wc-VF clauses).

(14) (Ich weiß nicht,) welches Buch daß der Hans gekauft hat.
(I know not) which book that the Hans bought has
'I dont know which book Hans bought.'

Sentences like (14) interact with the Verum Focus data in an interesting way: Verum Focus can only be achieved by stressing *daß*, but not by stressing the w-constituent, *welches Buch*, mirroring the facts in V2 clauses, (13a), where stressing the constituent in the Vorfeld cannot result in Verum Focus, but stressing the verb in the V2 position can.

(15) a. (Jetzt möchte ich wissen) mit wem *daß* du getanzt hast.
Now I'd like to know with whom that you danced have
'Now, I'd like to know whith whom you did dance.'
b. # (Jetzt möchte ich wissen) mit *wem* daß du getanzt hast.

(15b) is, of course, a grammatical sentence in German, but the embedded clause cannot be interpreted as bearing Verum Focus.

3. Syntax

In this section, we will discuss our syntactic treatment of the major clause types of German, as exemplified by the following sentences.

(16) a. Daß jeder Student ein Buch liest.
that every student a book reads
'that every student reads a book.'
b. Liest jeder Student ein Buch?
reads every student a book
'Does every student read a book?'

c. Ein Buch liest jeder Student.
a book reads every student
'Every student reads a book.'

d. (Ich weiß nicht,) welches Buch jeder Student liest.
(I don't know) which book every student reads
'I don't know which book every student reads.'

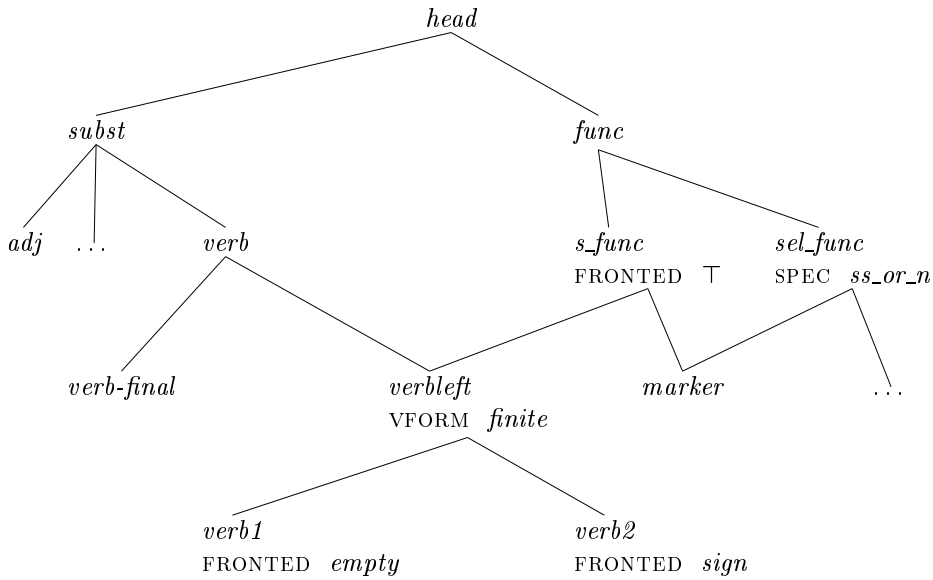
We adopt an analysis of German sentence structure along the lines of Richter 1997. In particular, we assume a binary branching structure, in which the complements are syntactically introduced according to their order on the valence lists. Since the resulting syntactic tree structure does not reflect word order, we combine it with a word order module that provides ordering principles on word order domains which may go beyond local trees (Reape 1992; Richter and Sailer 1995).[7]

We postulate a trace for topicalized and wh-moved constituents, and the NONLOCAL FEATURE PRINCIPLE of Pollard and Sag 1994. In contrast to Pollard 1996, we do not evoke a feature INV(erted) to treat verb movement. Furthermore, we do not use a HEAD FILLER SCHEMA to accommodate moved constituents to the syntactic structure of sentences. The most important change is an enrichment of the sort hierarchy subsumed by the sort *head* and the introduction of a new head feature, FRONTED, which, by virtue of attribute inheritance, is defined on markers and on some subsorts of *verb* and *func* as shown in figure 1. The value of this new attribute is either *empty* or a *sign*. Word order at the left periphery of finite sentences is governed by a word order principle:

(17) FRONTED ORDERING PRINCIPLE (FOP): In a finite sentence, S, with word order domain, D, the phonology of the unique word in

[7]We do not claim that a binary branching structure is the best way to treat German. We choose this analysis because (i), a binary branching structure simplifies the formulation of the SEMANTICS PRINCIPLE, and (ii), the valence attributes define an unambiguous syntactic position of the complements which avoids spurious ambiguities. Endorsing nonlocal word order domains, we avoid being forced to adopt a flat structure (Pollard 1996) or a word-order driven structure (Kiss 1995). Kiss 1998 motivates a word-order driven syntactic structure on the basis of order-dependent scopal behavior. Such an approach leads to spurious ambiguities for traces of non-scopal elements.
Throughout this study, we ignore adjuncts. We assume that adjuncts are syntactically realized higher than the complements, but, possibly, lower than the subject. To avoid spurious ambiguities, the hierarchical position of adjuncts might be determined by their relative surface word order. To avoid the problems mentioned above for complement traces, we reject adjunct traces. See Sailer 1997 for such an analysis, which takes into account the data collected in Hukari and Levine 1995.

FIGURE 1. The sort hierarchy for the sort *head*



D that bears the head feature FRONTED precedes the phonology of all other elements of D in the PHON value of S.

In effect, the FRONTED ORDERING PRINCIPLE says that, in each finite sentence, there must be exactly one word bearing a FRONTED feature, and that the phonology of this word comes first in the phonology of the overall sentence. The word order domain of a finite sentence contains its immediate constituents.[8] Since they do not bear the FRONTED attribute, words with head value *verb-final* do not qualify as the 'unique word' of (17) and are subject to word order principles concerning the position of the phonology of the verbal complex relative to the topology of its clause. Henceforth, we call the word of a finite sentence on which the head feature FRONTED is defined the *opener* of that sentence.

[8]For an exact specification of the pertinent relations and principles, see Richter 1997. The relational theory of word order presented there is modeled after the linearization theory of Richter and Sailer 1995. Our word order module crucially differs from Kathol 1995 in that we make explicit reference to the *word/phrase* status of the elements whose phonologies are to be ordered.

It is methodologically desirable to treat the binding of nonlocal dependencies by the same means, independent of whether a constituent is moved to the left periphery of a V2 clause or a w-VF clause. The HEAD-FILLER SCHEMA is however problematic if we want to treat the syntactic structure of c-VF and w-VF clauses uniformly. We therefore introduce a new principle to discharge slashed elements, even though we presuppose the NONLOCAL FEATURE PRINCIPLE.

(18) DEPENDENCY CANCELLATION PRINCIPLE (DCP): In a finite sentence, S, with word order domain, D, if D's opener has a FRONTED value of sort *sign*, then the TO-BIND SLASH value of the head-daughter is a singleton set whose element is token identical to the opener's FRONTED SYNSEM LOC value; otherwise, the TO-BIND SLASH value is the empty set.

Supposing the discussed changes to and additions of principles, all that is needed to account for the variation illustrated in (16) are the lexical entries for complementizers and finite verbs.

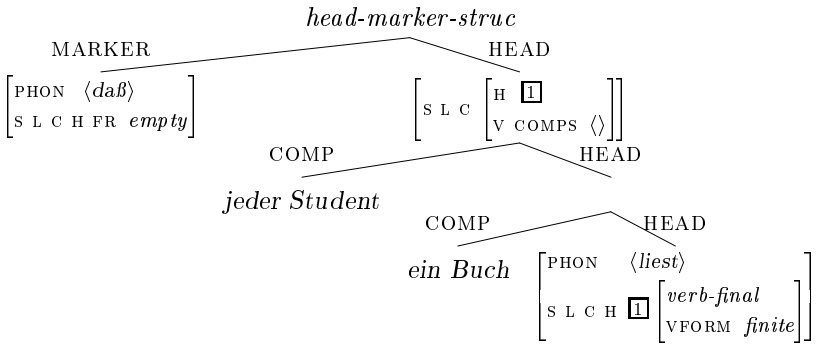
3.1. **Complementizers.** For c-VF clauses, we assume a marker *daß (that)* as shown in (19). It selects a verb with an empty COMPS list and a HEAD value of sort *verb*. With the FRONTED ORDERING PRINCIPLE it follows (by virtue of the sort hierarchy in figure 1) that, in every *head-marker-structure* with the marker *daß*, the head daughter must be a verb of sort *verb-final*. The word order of the marker also follows from the FOP. Figure 2[9] shows the syntactic structure of a c-VF clause.

(19) The marker *daß*:

$$\begin{bmatrix} \textit{word} \\ \text{PHON} \left\langle \textit{daß} \right\rangle \\ \text{S} \begin{bmatrix} \text{L C} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \textit{marker} \\ \text{FRONTED} \ \textit{empty} \end{bmatrix} \\ \text{SPEC L C} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \textit{verb} \\ \text{VFORM} \ \textit{finite} \end{bmatrix} \\ \text{VAL COMPS} \ \langle\rangle \end{bmatrix} \\ \text{MARKING} \ \textit{daß} \end{bmatrix} \\ \text{NONL} \begin{bmatrix} \text{INHER SLASH} \ \{\} \\ \text{TO-BIND SLASH} \ \{\} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

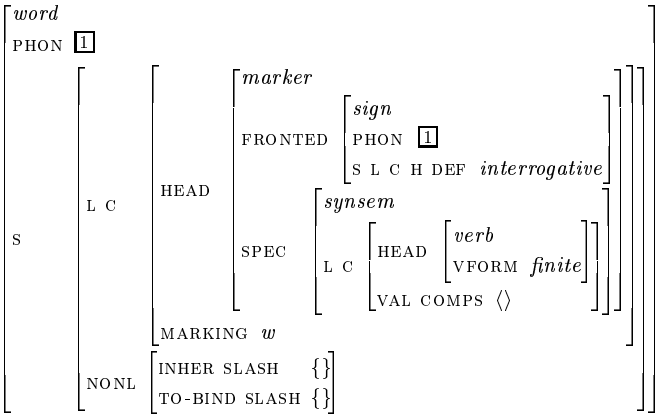
[9] To keep the descriptions simple, we use the following abbreviations for features: S(YNSEM), L(OCAL), NONL(OCAL), C(AT), H(EAD), FR(ONTED), V(AL), I(NHERITED), SL(ASH).

FIGURE 2. The structure of a c-VF clause: *daß jeder Student ein Buch liest*



For w-VF clauses, we introduce the marker shown in (20). Its MARKING value is *w*, and it selects a saturated finite verb projection with its SPEC attribute. In addition, it has a *sign*-valued FRONTED attribute. Finally, the PHON value of the marker is identical to the PHON value of its FRONTED constituent. These four properties guarantee that exactly one constituent moves out of the finite verbal projection selected by its SPEC value into the FRONTED value. Moreover, the phonology of the moved constituent appears in the overall phonology. As shown in figure 3, the DEPENDENCY CANCELLATION PRINCIPLE ensures that the LOCAL value of the sign in FRONTED is identical to the element in the TO-BIND SLASH value of the head-daughter in the sentence.

(20) The w-marker:

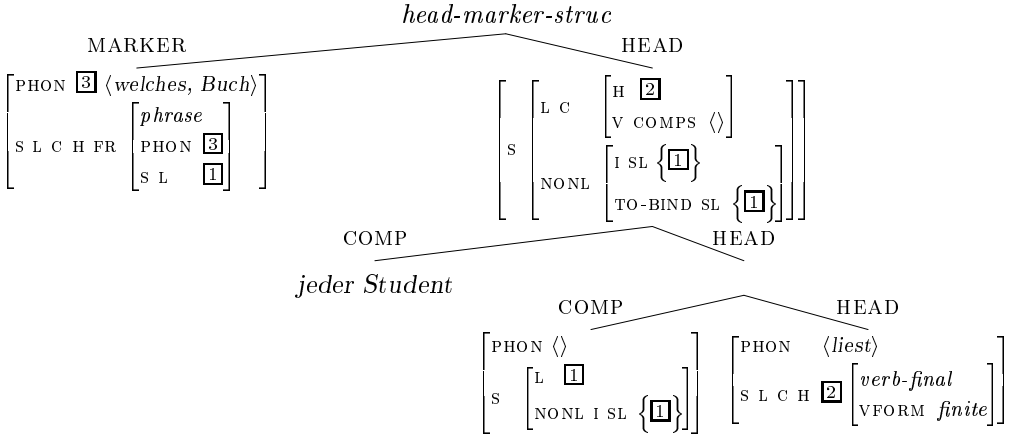


The w-marker imposes a restriction on the sign in FRONTED: It must be an interrogative w-constituent. As the FRONTED value is a sign, whether that sign is a w-constituent or not is available in any adequate analysis of pied piping. To our knowledge, the following options have been proposed: Pollard and Yoo 1998 proposes a feature percolation mechanism for the traditional nonlocal feature, QUE, specific for English, which would need to be modified for German, whereas Feldhaus 1999 introduces a head feature FORM and an amalgamation constraint to account for the corresponding German data. In (20), we adopt a proposal by Winhart 1997 that identifies w-constituents in German by the value *interrogative* of a head feature, DEF, which plays a prominent role in Winhart's comprehensive DP analysis. In particular, discussing criteria for functional categories established by Abney 1987, she concludes that prepositions in German have certain functional qualities that are reflected by their inheriting head values from their complement, including the DEF value Winhart (1997, p. 369). The w-marker, (20), thus licenses w-constituents as exemplified in (21) as its FRONTED value.

(21) a. (Ich weiß nicht,) [was] Hans kauft.
b. (Ich weiß nicht,) [welchen Hund] Karl gefüttert hat.
c. (Ich weiß nicht,) [in welchem Dorf] Karl wohnt.

For dialects that exhibit a co-occurrence of *daß* and a w-constituent at the left periphery as illustrated in (14), we assume that the w-marker differs slightly from its standard German counterpart, (20), by having an additional proper or idiosyncratic phonological contribution, *daß*. The

FIGURE 3. The structure of a w-VF clause: *welches Buch jeder Student liest*



only difference in the lexical entry concerns the description of the PHON value of the marker and is indicated in (22). The rest of the lexical entry is supposed to be identical with (20).

(22) $\begin{bmatrix}\text{PHON} & \boxed{1}\\ \text{S L C H} & \boxed{2}\end{bmatrix} \wedge \exists x$ `construct-phon`$(\boxed{1}, \boxed{2}, {}^{x}[\langle da\beta \rangle])$

As we will see immediately, the effect of the relation `construct-phon` in (22) is to concatenate the phonology of the w-constituent in FRONTED and the phonology $\langle da\beta \rangle$. The resulting phonological string is the phonology of the dialectal w-marker. In (14), this is *welches Buch daß*.

The lexical treatment of the left periphery enables us to reduce the difference between standard German w-VF clauses and dialectal wc-VF clauses to lexical variation. The simplicity of the account is also motivated, because, even for speakers of dialects, standard German w-VF clauses are grammatical, and because, for these speakers, the standard and the dialectal variant are distributionally free variants of each other.

3.2. **Finite Verbs.** We assume a lexical entry for finite verbs as illustrated in (23) for the verb *liest* (*reads*). Note the parallels of the lexical entry of a finite verb with the dialectal w-marker.

(23) Lexical entry of a finite verb (*liest*):

$$\left[\begin{array}{l}\textsc{phon}\ \boxed{1} \\ \\ \textsc{s}\quad\left[\begin{array}{ll}\textsc{l c} & \left[\begin{array}{ll}\textsc{head} & \boxed{2}\begin{bmatrix} verb \\ \textsc{vform}\ finite\end{bmatrix} \\ \textsc{val comps} & \langle \textsc{np}[nom],\ \textsc{np}[acc]\rangle \\ \textsc{marking} & unmarked\end{array}\right] \\ \textsc{nonl inher slash} & \{\}\end{array}\right]\end{array}\right]$$

$\wedge\ \exists x\ \texttt{construct-phon}(\boxed{1}, \boxed{2}, {}^{x}[\langle liest\rangle])$

The lexical entry for *liest* uses the relation `construct-phon` in (24).[10]

(24) The definition of `construct-phon`:

$$\begin{aligned}\forall x \forall y \forall z\ \texttt{construct-phon}(x, y, z)\ \leftrightarrow\ & \\ (\neg\ {}^{y}\left[\textsc{fronted}\ sign\right] \wedge x = z\ & \\ \vee\ \exists \boxed{1}\ \texttt{append}(\boxed{1}, z, x) \wedge\ {}^{y}\left[\textsc{fronted phon}\ \boxed{1}\right])\ & \end{aligned}$$

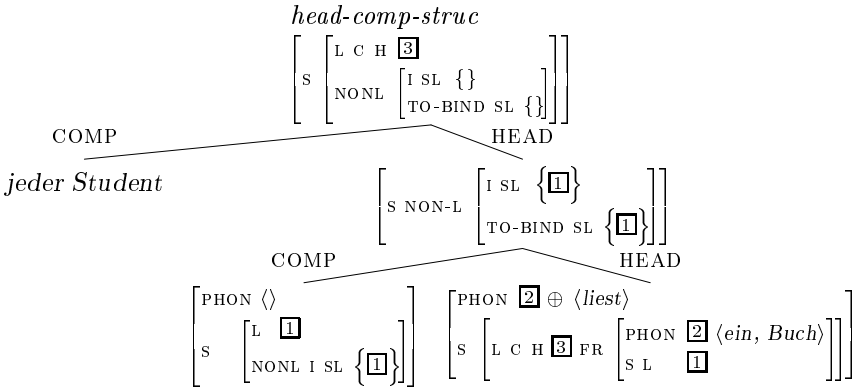
The lexical entries for finite verbs are underspecified with respect to the subsort of *verb*. Crucially, the PHON value of a finite verb in the model of the grammar is sensitive to the actual subsort of *verb*. This is expressed by the relation `construct-phon`. If a verb like *liest* has a HEAD value of sort *verb-final* or *verb1*, its PHON value simply consists of the string *liest*, since either FRONTED is not defined on its head or its value is not a *sign* (second line of 24). If its head is of sort *verb2*, the phonology of the FRONTED constituent is the initial part of the verb's PHON value (third line of 24).

We will first consider the structure of a V1 clause as in (16b). The finite verb, *liest*, in its word order domain has a HEAD value of sort *verb1* and is thus specified for FRONTED, and, due to the FOP, it is phonologically realized at the beginning of the sentence. Because its FRONTED value is *empty* (by virtue of the signature given in figure 1), the PHON value of the verb is simply *liest*.

In the verb second sentence (16c), the PHON value of the finite verb is *ein Buch liest*. In contrast to the w-marker, V2 verbs do not restrict their FRONTED constituents to w-constituents. By analogy to the treatment of the w-marker, there is a nonlocal dependency relation between the FRONTED value of the verb and the the sentence in which it occurs. This is enforced, once again, by the DEPENDENCY CANCELLATION PRINCIPLE

[10] To formalize relations in HPSG, we employ RSRL (Richter, Sailer, and Penn 1999) in the AVM notation of Richter 1999. RSRL extends King's 1989 SRL to allow for (bounded) quantification and relations and provides a model theory for HPSG grammars similar to the ones discussed in King 1999 and Pollard 1999.

FIGURE 4. The structure of a V2 clause: *ein Buch liest jeder Student*



given in (18). In the case of (16c), the first projection of the opener is the head-daughter of the sentence. Its TO-BIND SLASH element is, therefore, identical with the LOC value of the fronted constituent. In figure 4, the structure of sentence (16c) is shown.

In the present approach, the morphological variation exhibited by fronted verbs in, e.g., Dutch can be expressed directly in the lexicon, as the requisite information about both the subject and the Vorfeld constituent are lexically available to V2 verbs. The pertinent lexical entries have the following restrictions:

(25) a. *hebt*: if there is a FRONTED attribute, then the LOCAL values of the sign in FRONTED and that of the element in the SUBJ list are identical.

b. *heb*: there is a FRONTED feature, and if it is *sign*-valued, then the LOCAL values of the sign in FRONTED and that of the element in the SUBJ list are **not** identical.

As can be checked, these lexical specifications will account for the data in (1): *hebt* occurs in all VF clauses and only in V2 clauses with the clausemate subject as topicalized constituent. *heb* occurs only in *verbleft* clauses, i.e., V1 and V2 clauses, and only where the FRONTED value is not the clausemate subject, by being either *empty* or a different constituent.

The data on Verum Focus presented in section 2 can also be accounted for in a natural way. The head feature FRONTED on markers and finite

verbs in V1 and V2 clauses designates openers as a natural class of words. Verum Focus can, thus, be characterized by a generalization over words: In German, Verum Focus can occur if the opener of a finite sentence is focused. The opener is focused by stressing its idiosyncratic phonological contribution, if there is one, or otherwise by stressing the phonology of the w-pronoun of its FRONTED constituent. It is interesting to note that in w-VF sentences with a complex w-constituent like (13d), Verum Focus is not achieved by stressing the focus exponent of the phrase, which would be *Buch*, but rather by stressing the w-phrase, *wessen*. Stress on *Buch* results in a very different and unrelated effect. In other words, there is a clear distinction between stressing the phonetically invisible w-marker by way of its w-constituent FRONTED value and stressing the w-constituent itself.

For the dialectal data in (15), our generalization on Verum Focus predicts the data correctly: Because the dialectal w-complementizer in (22) has an idiosyncratic phonological contribution, *daß*, it is this contribution that receives stress in order to focus the opener. This corresponds to V2 clauses, where it is stressing the idiosyncratic phonological contribution of the finite verb that may trigger Verum Focus.

4. Semantics

We supplement the syntactic analysis of this study with a semantic treatment of sentences which combines the insights of the semantic approach to questions put forward by Groenendijk and Stokhof 1982 with techniques of underspecified semantics (Reyle 1993; Copestake, Flickinger, Malouf, Riehemann, and Sag 1995; Egg 1998). We will show that the opener not only plays an important role in the syntax of German clauses but that it also is the element that contributes the main functional aspects to the logical form of a clause.

We assume that the logical form of a sentence is an HPSG representation of an expression of Ty2 (Gallin 1975). Ty2 expressions as logical forms in HPSG are defined as objects in the denotation of a sort *meaningful-expression* (*me*) by Richter and Sailer 1997 and Richter and Sailer 1999a. They define the sort hierarchy below *me*, the appropriateness conditions for *me* and its subsorts, and the principles that ensure a proper correspondence between (equivalence classes of) *me* objects and expressions of Ty2 in the denotation of a grammar. To give a concrete example, we take (27) to be the logical form of (26). w is a world index.

(26) Welcher Student liest (mindestens) ein Buch?
which student reads at least one book

(27) $\lambda p.\exists x[\mathsf{student'}_w(x) \wedge p = \lambda w.\exists y[\mathsf{book'}_w(y) \wedge \mathsf{read'}_w(x,y)]]$

There are several possibilities to derive (27). Richter and Sailer 1999b and Richter and Sailer 1999a explore a version of semantic composition that only allows functional application (and β-reduction) at phrasal nodes. Scope ambiguities are handled by type-shifting operations at word level, following ideas by Hendriks 1993. In Richter and Sailer 1997, the logical form of a clause corresponds to an underspecified representation as defined by Bos 1995. The underspecified representations or 'unplugged' expressions of Ty2 denote a set of disambiguated specifications. The goal of the present analysis is to apply the techniques of underspecification to semantic composition, always deriving disambiguated logical forms like (27) at clause level. For that purpose, we structure the CONTENT of signs as follows:

(28) *The sort content*

content	TOP	*me*
	MAIN	*me*
	PARTS	*list(me)*

The TOP value of a clause is interpreted as its logical form. The MAIN value is needed for technical reasons only. The PARTS list contains the components of the overall logical form as contributed by the words that occur in the clause. The basic insight stemming from the literature on underspecified semantics is that the semantic contribution of a single lexical item can be discontinuous in the logical form of a clause. For example, in (27) the finite verb contributes its lexical meaning, $\mathsf{read'}_w(x,y)$, and the two parts of the interrogative operator, $\lambda p.\phi$ and $p = \lambda w.\psi$. As can be seen, the quantifier introduced be the w-word, $\exists x$, and the quantifier introduced by *ein*, $\exists y$, intervene between the contributions of the finite verb.

(29) *Well-formedness conditions for content*
In each *content*,
1. there are no multiple occurrences of the same element on the PARTS list,
2. the MAIN value is an element of the PARTS value,
3. all elements of the PARTS list are components of the TOP value.

We will first consider the logical form of the c-VF clause in (30). In (31), we list the MAIN values of the words that occur in that sentence.

For each of these words, the PARTS value is a list whose only element is identical to the MAIN value. The TOP value of the words need not be specified in the lexical entry, because, where relevant, it will ultimately be properly constrained by the SEMANTICS PRINCIPLE.

(30) a. daß jeder Student (mindestens) ein Buch liest.
that every student (at least) one book reads

b. $\exists w \forall x[\mathsf{student'}_w(x) \rightarrow [\exists y \mathsf{book'}_w(y) \wedge \mathsf{read'}_w(x,y)]]$

(31) a. *liest*: $\boxed{1}\mathsf{read'}_w(x,y)$

b. *Buch*: $\boxed{2}\mathsf{book'}_w(y)$

c. *ein*: $\boxed{4}\exists y[\alpha \wedge \beta]$

d. *Student*: $\boxed{3}\mathsf{student'}_w(x)$

e. *jeder*: $\boxed{5}\forall x[\gamma \rightarrow \delta]$

f. *daß*: $\boxed{6}\exists w[\epsilon]$

In our abbreviatory Ty2 notation for descriptions of *me* objects, we use Greek letters to indicate parts of the *me* object that are not exhaustively described. For example, the expression in (31c) is an abbreviation for the following description:

(32)
$$\begin{bmatrix} \textit{existential} & \\ \text{TYPE} & \textit{truth} \\ \text{VAR} & \textit{var} \\ \text{SCOPE} & \begin{bmatrix} \textit{conjunction} & \\ \text{TYPE} & \textit{truth} \\ \text{ARG1} & [\textit{me}] \\ \text{ARG2} & [\textit{me}] \end{bmatrix} \end{bmatrix}$$

The semantic composition in clauses is governed by the SEMANTICS PRINCIPLE (SP). It is modeled after the SP given for unplugged Ty2 in Richter and Sailer 1997. The version given in (33) is, however, simpler because here we are not interested in truely underspecified representations of clauses. We leave out the conditions for *head-adjunct-phrase*s.

(33) SEMANTICS PRINCIPLE
In each *headed-phrase*,
1. the TOP value of the head and the mother are identical,
2. the MAIN value of the head and the mother are identical,[11]

[11] For simplicity, we take determiners to be the head of quantified NPs.

3. the PARTS value contains exactly all elements of the PARTS values of the daughters,
4. the following conditions hold:
   (a) if the head is a quantifier with a MAIN value of the form $Qx[\rho \circ \nu]$[12], then the MAIN value of the non-head is a component of ρ,
   (b) if the non-head is a quantified NP with a MAIN value of the form $Qx[\rho \circ \nu]$, then the MAIN value of the head is a component of ν, and
   (c) in a *head-marker-phrase*, the MAIN value of the non-head is identical to the TOP value of the mother.

The first two clauses of the SP determine the values of the TOP and the MAIN attributes along the head path of a phrase. The third clause states that the PARTS values of the daughters are collected in the mother's PARTS value. The fourth clause regulates the relative embedding of the elements of the mother's PARTS list. We only list the three cases that are needed for our discussion below. For a broader coverage, further conditions must be added. Let us illustrate the application of the SP with the derivation of the logical form of sentence (30).

In the CONTENT value of the NP *ein Buch*, the MAIN value is identical to the one of the determiner by the second clause of the SP. All elements of the PARTS lists, (31b) and (31c), are collected in the NP's PARTS list. The MAIN value of the quantifier is of the form $\exists y[\alpha \wedge \beta]$. Therefore, due to clause (4a) of the SP, it is required that the MAIN value of the noun *Buch*, $\boxed{2}$, be a component of α. We use an infix notation for the binary relation `component` ($\lhd$). Two objects, o and o', are in the relation, $\lhd$, exactly if o is identical to o', or an attribute, α, is defined on o' that leads to an object o'', and o and o'' are in $\lhd$.

(34) *ein Buch*:
$$\begin{bmatrix} \text{TOP} & me \\ \text{MAIN} & \boxed{4} \\ \text{PARTS} & \langle \boxed{4}, \boxed{2} \rangle \end{bmatrix} \wedge \boxed{2} \lhd \alpha$$

The derivation of the logical form of the NP *jeder Student* is parallel and yields (35).

---

[12] $Qx[\rho \circ \nu]$ is shorthand for the description $\begin{bmatrix} quantifier & \\ \text{VAR} & var \\ \text{SCOPE} & \begin{bmatrix} l\text{-}const & \\ \text{ARG1} & \rho \\ \text{ARG2} & \nu \end{bmatrix} \end{bmatrix}$.

(35) *jeder Student*:

$$\begin{bmatrix} \textsc{top} & me \\ \textsc{main} & \boxed{5} \\ \textsc{parts} & \langle \boxed{5}, \boxed{3} \rangle \end{bmatrix} \wedge \boxed{3} \lhd \gamma$$

Next, we consider the logical form of the VP. In that case, clause (4b) of the SP is relevant. It has the effect of bringing the MAIN value of the finite verb, $\boxed{1}$, into the nuclear scope of the quantifier. This is expressed by the relational formula $\boxed{1} \lhd \beta$, where β is the nuclear scope of *ein* as given in (31c).

(36) *ein Buch liest*:

$$\begin{bmatrix} \textsc{top} & \boxed{6} \\ \textsc{main} & \boxed{1} \\ \textsc{parts} & \langle \boxed{1}, \boxed{4}, \boxed{2} \rangle \end{bmatrix} \wedge \boxed{1} \lhd \beta$$

The subject NP can be combined with the VP in exactly the same way, as indicated in (37).

(37) *jeder Student ein Buch liest*:

$$\begin{bmatrix} \textsc{top} & \boxed{6} \\ \textsc{main} & \boxed{1} \\ \textsc{parts} & \langle \boxed{1}, \boxed{4}, \boxed{2}, \boxed{5}, \boxed{3} \rangle \end{bmatrix} \wedge \boxed{1} \lhd \delta$$

The semantics of the sentence in (30) can now be derived by considering its highest local tree, which is a *head-marker-phrase*. Clause (4c) of the SP requires that the overall TOP value be identical to the TOP value of the marker daughter. That leads to an existential quantification over the world index, w.

(38) *daß jeder Student ein Buch liest*:

$$\begin{bmatrix} \textsc{top} & \boxed{6} \\ \textsc{main} & \boxed{1} \\ \textsc{parts} & \langle \boxed{1}, \boxed{4}, \boxed{2}, \boxed{5}, \boxed{3}, \boxed{6} \rangle \end{bmatrix}$$

In our example, the semantic contribution of the marker, *daß*, is an existential quantification over the world index. We regard that quantification as the functional content of a declarative sentence. For interrogative questions, we follow the semantic analysis of Groenendijk and Stokhof 1982 and von Stechow 1993, and assume the following logical form for an embedded interrogative:

(39) a. ... ob jeder Student (mindestens) ein Buch liest.
whether every student (at least) one book reads

b. $\lambda p.p = \lambda w.\forall x[\mathsf{student'}_w(x) \rightarrow \exists y[\mathsf{book'}_w(y) \wedge \mathsf{read'}_w(x, y)]]$

In the case of an interrogative clause, the functional content is a lambda abstraction over propositions, i.e., semantic entities of type $\langle s,t\rangle$. The treatment of w-interrogatives in the present semantic framework presupposes that the functional content in question consists of two components, the abstraction, $\lambda p.\phi$, and the equation, $p = \lambda w.\psi$. We assume that these components are elements of the PARTS values of the interrogative complementizer *ob*, of the w-marker, and of fronted w-constituents in both w-VF clauses and w-V2 clauses. In (40), the components of functional contents are summarized:

(40)

Declarative clauses:	$\exists w\phi$
Interrogative clauses:	$\lambda p.\phi$, $\boxed{1}p = \lambda w.\psi$, $\boxed{1} \lhd \phi$

In (41), a description of the CONTENT value of *ob* is given as it appears in the lexical entry.

(41)
$$\begin{bmatrix} content & \\ \text{TOP} & me \\ \text{MAIN} & \boxed{6} \\ \text{PARTS} & \langle \boxed{6}\lambda p.\phi, \boxed{7}p = \lambda w.\psi \rangle \end{bmatrix} \wedge \boxed{7} \lhd \phi$$

With our CONTENT descriptions being part of the lexical entries of the interrogative marker and of the other words in sentence (39a), respectively, the SP leads to the logical form (39b) for (39a). The derivation proceeds parallel to sentence (30).

Before we consider the semantic contribution of fronted verbs, it should be noted that none of our principles prevents the overall TOP value from containing semantic material that is not explicitly described in the lexical entries of the words. This is merely a technical observation. One possible solution is sketched in the appendix (section 8). It involves a slight extension of the signature and an additional principle, but leaves the SP and our fundamental assumptions unchanged. For the time being, we proceed with the simplified version.

We assume that the functional contents for declarative and interrogative markers shown in (40) can also be part of finite verbs that bear the attribute FRONTED. For the standard case in which the same verb form can be used in both fronted and non-fronted position, a relation, `construct-sem`, defines the possible CONTENT values of a finite verb in a single lexical entry. Parallel to what the relation `construct-phon` does for the phonology, `construct-sem` makes the CONTENT of a verb dependent on its actual FRONTED value.

(42) a. The semantic part of the lexical entries for finite verbs:

$$\begin{bmatrix} \textit{word} \\ \textsc{syns loc} \begin{bmatrix} \textsc{cat head } \boxed{1} \\ \textsc{cont } \boxed{3} \begin{bmatrix} \textit{cont} \\ \textsc{main } \boxed{2}\ \mathsf{read'}_w(x,y) \end{bmatrix} \end{bmatrix} \end{bmatrix}$$
$$\wedge\ \texttt{construct-sem}(\boxed{1},\boxed{2},\boxed{3})$$

b. $\forall x \forall y \forall z\ \texttt{construct} - \texttt{sem}(x,y,z) \leftrightarrow$

$$\left(\neg\ {}^{x}\left[\textsc{fr}\ \textit{empty} \vee \textit{sign}\right] \wedge\ {}^{z}\begin{bmatrix} \textit{cont} \\ \textsc{parts}\ \langle y \rangle \end{bmatrix}\right) \vee$$
$$\left({}^{x}\left[\textsc{fr}\ \textit{empty} \vee \textit{sign}\right] \wedge\ {}^{z}\begin{bmatrix} \textit{cont} \\ \textsc{top}\ \ \boxed{1}\ \exists w.\phi \\ \textsc{parts}\ \langle \boxed{1}, y \rangle \end{bmatrix} \wedge\ y \lhd \phi\right) \vee$$
$$\left({}^{x}\left[\textsc{fr}\ \textit{empty} \vee \textit{sign}\right] \wedge\ {}^{z}\begin{bmatrix} \textit{cont} \\ \textsc{top}\ \ \boxed{1}\ \lambda p.\phi \\ \textsc{parts}\ \langle \boxed{1}, \boxed{2} p = \lambda w.\psi, y \rangle \end{bmatrix} \wedge\ \boxed{2} \lhd \phi \wedge\ y \lhd \psi\right)$$

The first disjunct in the definition of the relation `construct-sem` treats the case of non-fronted finite verbs. In the second disjunct, the declarative functional content is an element of the PARTS list of the verb and the value of TOP. Finally, in the last disjunct, the interrogative content components are part of the verb's CONTENT. Let us now consider the derivation of the logical form of a V2 declarative clause.

(43) a. Jeder Student liest (mindestens) ein Buch.

b. $\exists w \forall x[\mathsf{student'}_w(x) \rightarrow [\exists y \mathsf{book'}_w(y) \wedge \mathsf{read'}_w(x,y)]]$

For the words that occur in the clause, their logical form has already been specified. For the derivation of the logical form of the sentence, the CONTENT of the fronted finite verb (with phonology *jeder Student liest*) and of the trace of the moved constituent are of particular interest. As indicated in (44), the relation `construct-sem` defines the declarative semantic component as the verb's TOP value, since it has a non-interrogative constituent as its FRONTED value.

(44) *jeder Student liest*:

$$\begin{bmatrix} \textsc{top} & \boxed{6} \\ \textsc{main} & \boxed{1} \\ \textsc{parts} & \langle \boxed{6} \exists w.[\epsilon], \boxed{1} \mathsf{read'}_w(x,y) \rangle \end{bmatrix} \wedge\ \boxed{1} \lhd \epsilon$$

(45) *t*:

$$\begin{bmatrix} \textsc{top} & \textit{me} \\ \textsc{main} & \boxed{5} \\ \textsc{parts} & \langle \boxed{5}, \boxed{3} \rangle \end{bmatrix} \wedge\ \boxed{3} \lhd \gamma$$

Notice that the semantic contribution of the trace, (45), is identical to the logical form of *jeder Student*. With the logical forms (44) and (45), the correct logical form for V2 declarative clauses can be derived with the SP. The semantic contribution of the constituent in the Vorfeld enters the overall semantics through its trace.

For V1 interrogatives, we assume a logical form that is identical to that of *ob* interrogatives. An example is given in (46).

(46) a. Liest jeder Student (mindestens) ein Buch?
reads every student (at least) one book

b. $\lambda p.p = \lambda w.\forall x \mathsf{student'}_w(x) \rightarrow [\exists y \mathsf{book'}_w(y) \wedge \mathsf{read'}_w(x,y)]$

The semantic derivation of (46a) follows from the CONTENT value of the finite verb in the way we have seen above.

(47) *liest*:

$$\begin{bmatrix} \text{TOP} & \boxed{7} \\ \text{MAIN} & \boxed{1} \\ \text{PARTS} & \langle \boxed{7}\lambda p.\phi, \boxed{8}p = \lambda w.\psi, \boxed{1} \rangle \end{bmatrix} \wedge \boxed{1} \lhd \psi \wedge \boxed{8} \lhd \phi$$

Finally, we turn to V2 interrogative clauses. Structurally identical to V2 declaratives, they have a w-phrase in the FRONTED value of the opener. (48) shows an example sentence together with its logical form.

(48) a. Welcher Student liest (mindestens) ein Buch?
which student reads (at least) one book

b. $\lambda p.\exists x \mathsf{student'}_w(x) \wedge p = \lambda w.[\exists y \mathsf{book'}_w(y) \wedge \mathsf{read'}_w(x,y)]$

Our principles allow that several words contribute identical parts to the logical form of a phrase. In the case of w-V2 clauses, both the fronted w-phrase (and, therefore, its trace) and the finite verb contain interrogative content components on their PARTS lists. The CONTENT value of the finite verb is, thus, identical to the verbal CONTENT in V1 interrogatives. The CONTENT value of the w-word *welcher* is given in (49).

(49) *welcher*:

$$\begin{bmatrix} \text{TOP} & me \\ \text{MAIN} & \boxed{11} \\ \text{PARTS} & \langle \boxed{9}\lambda p.\phi, \boxed{11}\exists x \gamma \wedge \delta, \boxed{12}p = \lambda w.\psi \rangle \end{bmatrix} \wedge \boxed{11} \lhd \phi \wedge \boxed{12} \lhd \delta$$

Combining (49) with the CONTENT of *Student*, (31d), we arrive at the w-phrase's CONTENT:

(50) *welcher Student*:

$$\begin{bmatrix} \textsc{top} & \textit{me} \\ \textsc{main} & \boxed{11} \\ \textsc{parts} & \langle \boxed{9}, \boxed{11}, \boxed{12}, \boxed{3} \rangle \end{bmatrix} \wedge \boxed{3} \lhd \gamma$$

For convenience, we state the CONTENT value of the fronted finite verb in (51). It is important to note that the interrogative meaning components, 9 and 12, occur in the PARTS lists of both the w-phrase and the finite verb.

(51) *welcher Student liest*:

$$\begin{bmatrix} \textsc{top} & \boxed{9} \\ \textsc{main} & \boxed{1}\ \mathsf{read'}_w(x,y) \\ \textsc{parts} & \langle \boxed{9}\lambda p.\phi, \boxed{12}p = \lambda w.\psi, \boxed{1} \rangle \end{bmatrix} \wedge \boxed{1} \lhd \psi \ \wedge \ \boxed{12} \lhd \phi$$

The logical form of the finite verb regularly combines with the logical form of the direct object (known from (34)):

(52) *ein Buch*+V:

$$\begin{bmatrix} \textsc{top} & \boxed{9} \\ \textsc{main} & \boxed{1} \\ \textsc{parts} & \langle \boxed{9}, \boxed{12}, \boxed{1}, \boxed{4}, \boxed{2} \rangle \end{bmatrix} \wedge \boxed{1} \lhd \beta$$

The logical form of the w-phrase enters the clausal semantics through the trace. In our example, the trace is the non-head daughter of highest projection of the finite verb. (53) contains a description of the CONTENT of the clause as entailed by the SP. It can be checked that its TOP value corresponds to the logical form in (48b).

(53) t + [*ein Buch*+V]:

$$\begin{bmatrix} \textsc{top} & \boxed{9} \\ \textsc{main} & \boxed{1} \\ \textsc{parts} & \langle \boxed{9}, \boxed{12}, \boxed{1}, \boxed{4}, \boxed{2}, \boxed{11}, \boxed{3} \rangle \end{bmatrix} \wedge \boxed{1} \lhd \delta$$

w-VF interrogatives as in (54) do not introduce any new aspects to our analysis. The CONTENT of the w-marker looks like the CONTENT of *ob* in (41), but allows w-constituents to intervene between the two parts of the interrogative semantics. As an illustration, (55) depicts the CONTENT of the w-marker as it occurs in sentence (54).

(54) a. . . . , welcher Student (daß) ein Buch liest.
which student (that) a book reads

b. $\lambda p.\exists x[\mathsf{student'}_w(x) \wedge p = \lambda w.\exists y[\mathsf{book'}_w(y) \wedge \mathsf{read'}_w(x,y)]]$

(55) w-marker (*welcher Student (daß)*):

$$\begin{bmatrix} \textsc{top} & me \\ \textsc{main} & \boxed{9} \\ \textsc{parts} & \langle \boxed{9}\lambda p.\phi, \boxed{12}p = \lambda w.\psi \rangle \end{bmatrix} \wedge \boxed{12} \lhd \phi$$

In regard to questions, we assume that in w-V2 clauses, the fronted verb and the fronted w-constituent and, in w-VF clauses, the w-marker and the fronted w-constituent each contribute an interrogative operator. It must, then, be assured that they contribute identical interrogative meaning components. It is a welcome consequence of the fact that we presuppose a typed semantic representation language that there cannot be a second interrogative operator in a clause, unless it is a semantic argument of some functor. If, in the derivation of (48), the w-phrase contributed an interrogative operator, $\lambda q.\phi$, which was different from $\boxed{9}$, no possible logical form for the overall sentence existed according to our principles.

In section 3, we have seen that openers play a central role in the syntax of German finite clauses. The Opener Principle, (56), highlights the importance of the opener for the semantics of a German clause.

(56) Opener Principle
In every clause, the TOP value of the clause is an element of the PARTS list of its opener.

Since German normally does not allow in situ w-interrogatives, we must assume a principle similar to Rizzi's *wh-criterion*. For every interrogative clause, if there is a fronted constituent, then that constituent must also contain an interrogative operator. This principle is given in informal terms in (57).

(57) W-Criterion
In every clause, if the TOP value is of the form $\lambda p.\phi$, then if the FRONTED value of the opener is a sign, s, then the TOP value of the clause must be an element of the PARTS value of s.

The W-Criterion enables us to exclude in situ questions and to predict the correct readings for multiple w-questions.

(58) * Peter liest welches Buch?
Peter reads which book

If (58) is interpreted as a question, the finite verb, by virtue of being the opener of the clause, must contribute an interrogative meaning component. Then the W-Criterion requires that the interrogative meaning be present in the Vorfeld constituent as well. The proper name *Peter*,

however, does not have an interrogative contribution, and the sentence is ungrammatical (under a normal interrogative interpretation).

The W-CRITERION does not force the opener to have an interrogative semantic contribution, if there is a fronted interrogative constituent. The fact that it must have one follows from the OPENER PRINCIPLE. Assume that the opener is not interrogative but contains an interrogative constituent as its FRONTED value. If the opener is not interrogative, it can only be declarative in our small fragment grammar. It follows from the OPENER PRINCIPLE that the entire clause cannot be interrogative; it must be declarative. Thus, there is no place in the logical form of the clause for the interrogative operator contributed by the fronted constituent.

Example (59) is an instance of a multiple w-question. Ignoring details of its logical form, it can be assumed that there is only one interrogative operator in the clause. The W-CRITERION and the OPENER PRINCIPLE ensure that this operator is part of the CONTENT of the fronted verb and the Vorfeld constituent. Due to the typed representation language, the interrogative operator contributed by the direct object must be identical to those of the other two interrogative elements.

(59) a. Welcher Student liest welches Buch?
which student reads which book

b. $\lambda p.\exists x \mathsf{student}'_{w'}(x) \wedge \exists y \mathsf{book}'_{w'}(y) \wedge p = \lambda w.\mathsf{read}'_{w}(x, y)$

In complex sentences, which we did not address, there exists the possibility that more than one interrogative clause may occur in a sign.

(60) Wer weiß, welcher Student welches Buch liest?
who knows which student which book reads

In (60), there are three w-constituents and two interrogative clauses. Both subject w-constituents have moved into the FRONTED value of the opener of their respective clause. They are integrated in the finite verb in the case of the matrix clause and in the w-marker in the embedded interrogative. The W-CRITERION guarantees that the interrogative meaning of these w-constituents is interpreted in the clause to which they belong. The W-CRITERION is, however, neutral with respect to the interpretation of the w-constituent in the Mittelfeld of the embedded clause. Its interrogative meaning contribution can be identical to either the one of the embedded clause or the one of the matrix clause. This prediction is empirically correct (Feldhaus 1997; Pollard and Yoo 1998; Ginzburg and Sag 1998).

5. Comparison: German Sentence Structure in HPSG

Among the HPSG analyses of German sentence structure (Pollard 1996; Frank 1994; Kiss 1995; Müller 1999; Feldhaus 1999), the work of Kathol 1995 is the one that is most closely related to our proposal, because it takes topological fields and word order domains as fundamental notions of grammar. For reasons of space, our comparison will mainly focus on Kathol 1995. In section 5.2, we will briefly address some aspects of the other HPSG analyses.

5.1. **Topological Fields in HPSG.** Kathol 1995 and Kathol 1996 incorporate a variant of the traditional theory of topological fields (Engel 1970; Höhle 1986) directly into HPSG. Kathol assigns topological fields (TOPO values) to elements in the linearization domain of a clause, where the linearization domain essentially contains the immediate constituents of the clause. He captures the complementary distribution of fronted verbs and complementizers by a uniqueness requirement on certain members of linearization domains of sentences: There may be no more than one element bearing the TOPO value *cf* (complementizer-field), which characterizes finite verbs and complementizers at the left periphery. Kathol, nevertheless, has a filler-head constellation for fronted constituents (in the case of long extraction), and the TOPO value of fronted constituents can be either *vf* (Vorfeld) or *cf*.

Kathol's analysis covers the basic data. However, as he splits the left periphery into two independent topological fields, *vf* and *cf*, all data where the element in one field is dependent on the element in the other appear problematic. Firstly, as finite verbs do not have access to the Vorfeld constituent at the word level, the morphological variation discussed in section 2 cannot be accounted for at this level. One possibility to remedy this deficiency is to impose a principle on clauses that refers to both the element in the Vorfeld and the inflection paradigm of the finite verb. As such a principle seems to conflate the levels of syntax and morphology in an undesirable way, this seems to be rather unattractive.

A second problem with splitting up the left periphery arises with data such as (14). If a fronted w-constituent co-occurs with *daß*, both left-peripheral fields, *vf* and *cf*, are filled. Kathol (1995, p. 169 and p. 173) treats this distinction as a matter of parametric variation between standard German and Southern dialects, where the latter do not force the phonology of a marker to appear clause initially. The situations in which a complementizer can co-occur with a fronted constituent are, however, far more restricted than such a general parameter would suggest. In German

dialects, *daß* and the relativizer *wo* are the most common examples. For other complementizers, such as *ob* (*whether*), *wenn* (*if*) or *bevor* (*before*), there is great variation among speakers as to whether they accept them in combination with a fronted constituent or not. To affirm the availability of a pre-complementizer constituent in German dialects, Kathol refers to the two constructions illustrated in (61).

(61) a. [An Mantl daß da Xaver kaffd hod] hod neamt glaubt.
a coat that the Xaver bought has has no-one believed
'No-one believed that Xaver bought a coat.'

b. [Das Bier wenn ich noch trink] bin ich gleich besoffen.
the beer if I still drink am I soon drunk
'If I drink the beer, I'll soon be drunk.'

Since it is not the case, however, that all speakers who accept the wc–VF construction also accept the constructions in (61) the parametric account is too general. In addition, the constructions in (61) differ from the wc-VF construction in a significant way: They may only occur in the Vorfeld of a V2 clause. This is a restriction that does not apply to wc-VF clauses. It is, thus, far from clear whether a unified analysis as envisaged by Kathol is adequate.

Although Kathol offers a solution to the kind of data under discussion by constraining the DOM lists of clauses, verbal morphology and special lexical elements like *daß* in wc-VF constructions lead us to consider a lexicalized analysis to be more adequate; it also seems to be more in the spirit of the lexicalist orientation of many HPSG grammars.

5.2. **Other HPSG approaches.** All HPSG approaches to the German left periphery basically build on the assumption of a sentence structure as indicated in (62). Speaking in GB terms, the fronted constituent c-commands the rest of the clause. There is no selection between the fronted finite verb or the complementizer and the fronted constituent, apart from long distance SLASH specifications.

(62) a. [Ein Buch [liest jeder Student]]

b. [welches Buch [(*e*/daß) [jeder Student liest]]]

Under this structural assumption, the data showing differing verb inflection on finite verbs depending on which constituent occupies the Vorfeld are unexpected and hard to account for, as are the complementizer coordination data discussed above.[13]

[13]As pointed out to us by Anke Feldhaus (p.c.), the inflection data can, in principle, be accounted for as restrictions on inflection paradigms built into ID schemata that

More technical points of criticism concern the feature geometry used in some of the alternative approaches. In Pollard 1996, perhaps due to the lack of a semantic analysis, the only purpose of a boolean valued head feature INVERTED is to account for the linear order of the finite verb. Its function does not go beyond a direct descriptive device to account for the V2 phenomenon. The FRONTED attribute and its value are used to determine the inflection paradigm and the semantic contribution, and it identifies the opener, which is the main functional element of the clause and subject to the central principles of our analysis.

Müller 1999 and Feldhaus 1999 employ more than one ID schema to account for constituent fronting. Müller 1999 distinguishes a schema for V2 clauses from a schema for w-VF clauses, and Feldhaus 1999 distinguishes clauses with fronted w-constituents from clauses where the fronted constituent is not a w-constituent. Our analysis shows that, in principle, a unified analysis of constituent fronting is possible across all clause types in German.[14]

6. A Comparison to the GB Architecture

Our analysis of the left periphery of finite clauses exhibits some striking correspondences to standard GB approaches. A *word* that has a FRONTED feature can be regarded as standing in C^0 position. The FRONTED value of such a *word* corresponds to Spec-CP position. FRONTED is *empty* in the case of an empty Spec-CP position, and FRONTED has a *sign* value if a constituent moves to Spec-CP. In that respect, the FRONTED feature as a functional property of words is reminiscent of three of Fukui's properties of functional heads. Fukui (1986, p.31) says that functional heads can have specifier positions, these specifier positions are noniterable, and the elements occupying the specifier position are moved from within the complement of the functional head.

On the other hand, a lexicalized analysis does not run into GB-specific problems such as accounting for the projection level of the element in

look similar to the HEAD-FILLER-SCHEMA. Note that this solution departs from the selectional perspective on inflection put forward in Pollard and Sag 1994 and requires specialized attributes for inflection paradigms.

[14]Pollard 1996 and Kiss 1995 do not provide an analysis of w-VF clauses, so it is unclear how their treatment of constituent fronting in V2 clauses can best be extended to w-VF clauses.

Spec-CP or C^0. The complementary distribution of finite verbs and complementizers at the left periphery in V1, V2, and c-VF clauses is accounted for by the FRONTED ORDERING PRINCIPLE. The problem of w-VF clauses is solved by lexically specifying a complementizer that shares its phonology with the phonology of its FRONTED constituent. The positional effect of GB's functional category, C, is achieved by a new feature FRONTED on certain lexical elements. The semantic effects associated with the functional category C are lexically incorporated into openers. Verb movement is accounted for by a word order principle that is sensitive to the presence or absence of the feature FRONTED on words.

The present approach leads directly to a generalization of the traditional model of topological fields of German sentences (Engel 1970; Höhle 1986). Traditionally, it specifies a unique pattern for each sentence type. At the same time, these patterns only differ with respect to the left periphery. In Höhle's terminology and ignoring the fields for coordinating particles and left dislocation, VF clauses begin with the C field, V1 clauses with the FINIT field, and V2 clauses with a K field and a FINIT field. Under the present analysis, the left periphery is uniformly occupied by openers that distinguish all three possibilities. As a consequence, all sentence types are treated uniformly, while, for each sentence, its sentence type can be succinctly identified by the properties of its opener alone.

7. CONCLUSION

We have presented a uniform lexicalized treatment of the left periphery of German sentences. The traditional topological fields characterizing the left periphery are subsumed under a class of lexical elements (which we call openers) that are distinguished by a new head feature. This head feature is responsible for the order of those words and may host an additional constituent whose properties are, thus, accessible to lexical restrictions such as morphological variation and Verum Focus imposed by the host word. Openers determine the sentence type and, by their functional content, the semantic type of a sentence. The apparent violation of the Doubly Filled Comps filter in w-*daß* constructions in substandard dialects of German is reduced to lexical variation, and the additional lexical entry needed to analyze these data patterns with finite verbs. Moreover, the parallel behavior of both V2 verbs and the dialectal w-marker with respect to Verum Focus and inflection is predicted as a consequence of their parallel structure. Comparing our theory to previous approaches in GB and HPSG, we have argued that the lexical approach captures generalizations that have usually presented a problem to analyses in both frameworks.

8. Appendix

Let us finally address the issue that none of our principles prevents the overall TOP value of a phrase from containing semantic material that is not explicitly described in the lexical entries of the words it comprises. To overcome this problem, we extend the signature slightly. We assume that, in addition to the appropriateness specifications given in Richter and Sailer 1999a, there is a boolean valued attribute LICENSED which is defined for the sort *meaningful-expression* (*me*).[15]

The Semantics Principle remains as stated in (33) above. Every *me* object that is a component of a phrase's CONTENT value must be a component of the CONTENT value of at least one word that is dominated by the phrase. Only words can license semantic contributions. Those parts of the logical form that are explicitly contributed by a word are marked as LICENSED *plus*; the others (the "holes" in the approach of Bos 1995 and Richter and Sailer 1997) are marked as LICENSED *minus*. (63) contains a description that corresponds to the description in (32) of the MAIN value of the indefinite article, *ein*, in its lexical entry.

(63)
$$\left[\begin{array}{ll} \textit{existential} & \\ \textsc{licensed} & \textit{plus} \\ \textsc{type} & \textit{truth} \\ \textsc{var} & \textit{var} \\ \textsc{scope} & \left[\begin{array}{ll} \textit{conjunction} & \\ \textsc{licensed} & \textit{plus} \\ \textsc{type} & \textit{truth} \\ \textsc{arg}1 & \left[\begin{array}{ll} \textit{me} & \\ \textsc{licensed} & \textit{minus} \end{array}\right] \\ \textsc{arg}2 & \left[\begin{array}{ll} \textit{me} & \\ \textsc{licensed} & \textit{minus} \end{array}\right] \end{array}\right] \end{array}\right]$$

We assume a principle which we might call the Principle of Full Interpretation (PFI). It states that the logical form of an utterance consists exactly of semantic components which are lexically licensed within the utterance. More formally, this means that for every *me* object x that is a component of the utterance's CONTENT TOP value, there is an *me* object y with LICENSED *plus* such that x and y differ at most (recursively) with respect to their LICENSED values. Conversely, for every *me* x object with LICENSED *plus*, there is a component y of the utterance's TOP value such that x and y differ at most (recursively) with respect to their LICENSED values. To formalize this principle, we define a binary relation,

[15] The model-theoretic interpretaion of *me* objects ignores LIC values.

$\cong$. Two *me* objects x and y stand in the relation $\cong$ exactly if x and y are of the same sort and of identical semantic type, and the corresponding values of *me*-valued attributes on x and y also stand in the relation, $\cong$. Thus, $\cong$ defines an identity up to LIC values.

The PFI is given in (64). Without argument, we follow Richter 1997 in assuming that utterances are modelled by objects of sort *unembedded-sign* (*u-sign*), which is a subsort of *sign*.

(64) PRINCIPLE OF FULL INTERPRETATION (PFI):
$\left[u\text{-}sign\ \right] \rightarrow \phi \wedge \psi$

where ϕ is

$\forall x \forall \boxed{1}$
$\left(\left[\begin{array}{l} u\text{-}sign \\ \text{SYNS L CONT TOP ME } \boxed{1} \end{array} \right] \wedge {}^{x}\left[me \right] \wedge x \triangleleft \boxed{1} \rightarrow \exists y \left({}^{y}\left[\begin{array}{ll} me & \\ \text{LIC} & plus \end{array} \right] \wedge x \cong y \right) \right)$

and ψ is

$\forall x \forall \boxed{1}$
$\left(\left[\begin{array}{l} u\text{-}sign \\ \text{SYNS L CONT TOP ME } \boxed{1} \end{array} \right] \wedge {}^{x}\left[\begin{array}{ll} me & \\ \text{LIC} & plus \end{array} \right] \rightarrow \exists y \left(y \triangleleft \boxed{1} \wedge x \cong y \right) \right)$

Finally, whereever the relation $\triangleleft$ is evoked in section 4 between to objects, o and o', it needs to be replaced by the requirement that there exists an o'' that stands in the $\cong$ relation with o, and o'' is a component of o'.

REFERENCES

Abney, S. P. (1987). *The English noun phrase in its sentential aspect.* Ph. D. thesis, MIT.

Bayer, J. (1984). COMP in Bavarian syntax. *The Linguistic Review 3*, 209–274.

Bos, J. (1995). Predicate logic unplugged. In *Proceedings of the 10th Amsterdam Colloquium.*

Copestake, D. Flickinger, R. Malouf, S. Riehemann, and I. Sag (1995). Translation using Mininimal Recursion Semantics. In *Proceedings of The Sixth International Conference on Theoretical and Methodological Issues in Machine Translation (TMI-95)*, Leuven.

Egg, M. (1998). *Wh*-questions in Underspecified Minimal Recursion Semantics. *Journal of Semantics 15*, 37–82.

Engel, U. (1970). Regeln zur Wortstellung. In U. Engel (Ed.), *Forschungsberichte des Instituts für deutsche Sprache*, Volume 5, pp. 7 – 148. o.O.

Feldhaus, A. (1997). Eine HPSG-Analyse ausgewählter Phänomene des deutschen w-Fragesatzes. Technical Report 27, Institute for Logic and Linguistics, Heidelberg.

Feldhaus, A. (1999). German wh-interrogatives in HPSG. This volume.

Frank, A. (1994). Verb Second by Lexical Rule or by Underspecification. Arbeitsberichte des SFB 340 43, Universität Stuttgart.

Fukui, N. (1986). *A theory of category projection and its applications*. Ph. D. thesis, MIT.

Gallin, D. (1975). *Intensional and Higher-Order Modal Logic*. North-Holland, Amsterdam.

Ginzburg, J. and I. Sag (1998, July). English interrogative constructions. Draft.

Groenendijk, J. and M. Stokhof (1982). Semantic analysis of wh-complements. *Linguistics and Philosophy 5*, 175 –233.

Hendriks, H. (1993). *Studied Flexibility*. ILLC Dissertation Series 1995-5. Institute for Logic, Language and Computation, Amsterdam.

Hinrichs, E., D. Meurers, F. Richter, M. Sailer, and H. Winhart (1997). Ein HPSG-Fragment des Deutschen, Teil 1: Theorie. Arbeitspapiere des SFB 340 95, Universität Tübingen.

Höhle, T. N. (1986). Der Begriff "Mittelfeld". Anmerkungen über die Theorie der topologischen Felder. In A. Schöne (Ed.), *Kontroversen alte und neue. Akten des 7. Internationalen Germanistenkongresses Göttingen 1985*, Volume 3, pp. 329–340.

Höhle, T. N. (1992). Über Verum-Fokus im Deutschen. In J. Jacobs (Ed.), *Informationsstruktur und Grammatik*, pp. 112–141. Westdeutscher Verlag, Opladen.

Höhle, T. N. (1997). Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. Arbeitspapiere des SFB 340 90, Universität Tübingen.

Hukari, T. E. and R. D. Levine (1995). Adjunct extraction. *Journal of Linguistics 31*, 195 – 226.

Kathol, A. (1995). *Linearization-Based German Syntax*. Ph. D. thesis, Ohio State University.

Kathol, A. (1996). The syntax of Wh-quantifier retrieval. In R. Jonkers, E. Kaan, and A. Wiegel (Eds.), *Language and Cognition*, Number 5 in Yearbook of the research group for Theoretical and Experimental Linguistics, pp. 101–114. University of Groningen.

King, P. J. (1989). *A logical Formalism for Head-Driven Phrase Structure Grammar*. Ph. D. thesis, University of Manchester.

King, P. J. (1999). Towards truth in HPSG. This volume.

Kiss, T. (1995). *Merkmale und Repräsentationen*. Opladen: Westdeutscher Verlag.

Kiss, T. (1998). Notes from the border: Scope and variable binding in German. Manuscript.

Müller, S. (1999). *Deutsche Syntax deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche*. Number 394 in Linguistische Arbeiten. Tübingen: Max Niemeyer Verlag. To appear.

Pollard, C. and I. A. Sag (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press.

Pollard, C. J. (1996). On head non-movement. In H. Bunt and A. van Horck (Eds.), *Discontinuous Constituency*, pp. 279–305. Mouton de Gruyter.

Pollard, C. J. (1999). Strong generative capacity in HPSG. In G. Webelhuth, J.-P. Koenig, and A. Kathol (Eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, Chapter 18, pp. 281–297. CSLI Publications.

Pollard, C. J. and E. J. Yoo (1998). A unified theory of scope for quantifiers and *wh*-phrases. *Journal of Linguistics 34*, 415–445.

Reape, M. (1992). *A Formal Theory of Word Order: A Case Study of West Germanic*. Ph. D. thesis, University of Edinburgh.

Reis, M. (1985). Satzeinleitende Strukturen im Deutschen. Über COMP, Haupt- und Nebensätze, W-Bewegung und die Doppelkopfanalyse. In W. Abraham (Ed.), *Erklärende Syntax des Deutschen*. Tübingen.

Reyle, U. (1993). Dealing with ambiguities by underspecification: Construction, representation and deduction. *Journal of Semantics 10*(2), 123 – 179.

Richter, F. (1997). Die Satzstruktur des Deutschen und die Behandlung langer Abhängigkeiten in einer Linearisierungsgrammatik. Formale Grundlagen und Implemenitierung in einem HPSG-Fragment. In: (Hinrichs, Meurers, Richter, Sailer, and Winhart 1997).

Richter, F. (1999). RSRL for HPSG. This volume.

Richter, F. and M. Sailer (1995). Remarks on Linearization. Reflections on the Treatment of LP-Rules in HPSG in a Typed Feature Logic. Master's thesis, Universität Tübingen.

Richter, F. and M. Sailer (1997). Underspecified semantics in HPSG. In H. Bunt, L. Kievit, R. Muskens, and M. Verlinden (Eds.), *Proceedings of the Second International Workshop on Computational Semantics, 8–10 January 1997, Tilburg*, pp. 234–246.

Richter, F. and M. Sailer (1999a). A lexicalist collocation analysis of sentential negation and negative concord in French. This volume.

Richter, F. and M. Sailer (1999b). LF Conditions on Expressions of Ty2: An HPSG Analysis of Negative Concord in Polish. In R. D. Borsley and A. Przepiórkowski (Eds.), *Slavic in HPSG*, pp. 247–282. CSLI Publications. In press.

Richter, F., M. Sailer, and G. Penn (1999). A formal interpretation of relations and quantification in HPSG. This volume.

Sailer, M. (1997). Adjunkte in einer linearisierungsgrammatik: Syntax, semantik und wortstellung. In E. Hinrichs, W. D. Meurers, F. Richter, M. Sailer, and H. Winhart (Eds.), *Ein HPSG-Fragment des Deutschen. Teil 1: Theorie*, Volume 95 of *Arbeitspapiere des SFB 340*, pp. 249–318.

von Stechow, A. (1993). Die Aufgaben der Syntax. In J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann (Eds.), *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*, Volume 1, pp. 1–88. Walter de Gruyter.

Winhart, H. (1997). Die Nominalphrase in einem HPSG-Fragment des Deutschen. In: (Hinrichs, Meurers, Richter, Sailer, and Winhart 1997).

Universität Tübingen, Seminar für Sprachwissenschaft, Abteilung Computerlinguistik
E-mail address: {fr,mf}@sfs.nphil.uni-tuebingen.de