A GPSG ACCOUNT OF VP FRONTING IN GERMAN

Mark Johnson

1. Introduction

In this paper I analyse the surface constituent structure of the German VP. I focus on two hypotheses:

(A) The main verb and its objects combine to form a constituent which then combines with the AUX to form the VP (as indicated by the bracketing in (1))

and

(B) The main verb and the AUX combine to form a constituent which then combines with the verb's objects to form the VP (as in (2)).

(1) Fritz sagte, dass er [[Helmut gesehen] hatte]
Fritz said, that he Helmut seen had
"Fritz said that he had seen Helmut"

(2) Fritz sagte, dass er [Helmut [gesehen hatte]]

It has been generally assumed that hypothesis (A) correctly represents the constituent structure of the VP (eg. Bierwisch (1965)). In this paper I present data, some of which is not explainable only in terms of hypothesis (A), some of which is not explainable only in terms of hypothesis (B). A theory which generates all the data presented here would have to incorporate both hypotheses (A) and (B)[1]. Finally I show that this is exactly what we expect if an analysis similar to that of Nerbonne (1982) is combined with the proposed analysis of auxiliaries by Uszkoreit (1982) following that of Gazdar et al (1981) for English.

---

[1] In a theory of grammar with a transformational component, the constituent bracketing in either (1) or (2) could be regarded as "underlying", and a rule of AUX movement or a "restructuring" rule generate the other constituent bracketing. In a theory without a transformational component such an option is not available, and a grammar for German written in this framework must incorporate both analyses (1) and (2) directly.

2. Modal Flip

In this section I consider a phenomenon I have dubbed "Modal Flip". In clauses with a finite form of the perfective auxiliary haben and a modal auxiliary verb, two orderings of the verbal elements within VP are possible, as in (3) and (4).

(3) Er sagte, dass er Helmut sehen gekonnt hätte
He said that he Helmut see-Inf could-PP had-Subj
"He said that he could have seen Helmut"

(4) Er sagte, dass er Helmut hätte sehen können
had-Subj see-Inf could-Inf

The phenomenon has been investigated by Kunsman (1975) and Johnson (1982) and is still lacking a comprehensive treatment. Judgements on examples of Modal Flip more complex than (4), even those made by native speakers of German who are linguists often differ, but the existance of the phenomena, or of the grammaticality of (4), is not in doubt.

For the purposes of the paper it suffices to note that in (4) hatte is interposed between the direct object Helmut and the main verb sehen. Now, hypothesis (A) predicts that Helmut gesehen is a single constituent; a sister of the auxiliary.

But this cannot be the case in (4), since the direct object Helmut and the main verb sehen are not contiguous; moreover since the auxiliary hätte is interposed between Helmut and sehen, it is not possible that Helmut and sehen form a constituent that is a sister of hätte.

Thus hypothesis (A) does not hold for sentences like (4). Indeed, it would predict incorrectly that sentences like (5 a) or (5 b) would be grammatical.

(5 a) * Er sagte, dass er hätte Helmut sehen können

(5 b) * Er sagte, dass er Helmut sehen hätte können

Hypothesis (B) does not claim that Helmut and sehen are a constituent, so the fact that they are not contiguous in (4) presents no problem for it.

Thus only hypothesis (B) can satifactorily explain the "Modal Flip" data.

3. Fronting

In German, the finite verb in a main clause is preceded by one single phrasal constituent, for example an NP as in (6), or a PP as in (7). The sentence is ungrammatical if the finite verb is preceded by two constituents, as in (8).

(6) Seinen Vater hat Fritz gesehen
his father has Fritz seen
"(It was) his father Fritz has seen"

(7) Im Garten hat Fritz seinen Vater gesehen
In=the garden has Fritz his father seen
"(It was) in the garden Fritz has seen his father"

(8) * Im Garten seinen Vater hat Fritz gesehen

In this paper I assume that this is due to Fronting, which places a phrasal constituent in front of the finite verb, the actual formulation of which is not given here[2]. I am interested in Fronting in as much as it gives us a convenient test of constituenthood. If a string of words can be placed in front of the finite verb of a grammatical sentence, this is good evidence that they form a constituent in that sentence.

In this section I examine what parts of the VP can be fronted, thus identifying the constituent structure within the VP. A sentence with the subject in initial position is shown in (9).

(9) Er hat seiner Tochter ein Märchen erzählen können
He has his daughter:DAT a tale:ACC tell be=able
"He has been able to tell his daughter a tale"

Both the accusative and dative objects of the verb can be separately fronted (10a,b), but not both together (10c).

(10 a) Seiner Tochter hat er ein Märchen erzählen können

(10 b) Ein Märchen hat er seiner Tochter erzählen können

(10 c) * Seiner Tochter ein Märchen hat er erzählen können

The main verb erzählen can be fronted, optionally with either or both of its objects[3].

(11 a) Erzählen hat er seiner Tochter ein Märchen können

(11 b) Ein Märchen erzählen hat er seiner Tochter können

(11 c) Seiner Tochter ein Märchen erzählen hat er können

The critical fact is that the main verb and its objects can be fronted without the auxiliary. This is evidence that the main verb and its objects form a constituent that does not contain the auxiliary: ie. hypothesis (A).

---

[2] See Uszkoreit (1982) for a possible formulation of Fronting in a GPSG framework.

[3] This phenomenon led Nerbonne (1982) to propose the analysis on which the work presented here is based. He did not consider the interaction of auxiliaries and objects in Fronting, however.

But the auxiliary können can also be fronted with the main verb, as in (12 a), although not on its own, as in (12 b).

(12 a) Erzählen können hat er seiner Tochter ein Märchen

(12 b) * Können hat er seiner Tochter ein Märchen erzählen

The verbal objects can also be added to the AUX plus main verb combination, as in (13a,b).

(13 a) Ein Märchen erzählen können hat er seiner Tochter

(13 b) Seiner Tochter ein Märchen erzählen können hat er

The critical fact in (12) is that the main verb and auxiliary can be fronted without any of the verb's objects. This is evidence that the main verb and the auxiliary form a constituent that does not contain the verb's objects; i.e., hypothesis (B).

If we wished to avoid the necessity of making hypothesis (B) part of our description of German we might change our explanation of the fronting process in such a way that it would generate (12 a) even though erzählen können is not a constituent. Just how this might be done is not clear, but even if an alternative account of fronting could be given, the "Modal Flip" data of the last section could not be explained.

4. A GPSG Account

In this section I show how a GPSG account of this data naturally predicts that a VP consisting of object NPs, a finite verb and auxiliaries will have two analyses, coresponding to hypothesis (A) and hypothesis (B). My GPSG account of the preceding data is a development of work done by Nerbonne. Where not otherwise specified, I am assuming a GPSG syntax conventions as in Gazdar and Pullum (1982), and a type driven semantics as in Klein and Sag (1981).

In this account, each verb's lexical entry has features identifying the categories[4] it subcategorizes for.

[4] These features identify not only the major category type but also subcategorization, in German this includes Case.

(14)

V [- NP_{dat}] [- NP_{acc}]	e.g. erzählen, geben etc. "to tell", "to give"
V [- NP_{acc}]	e.g. lesen etc. "to read"
V [- NP_{dat}]	e.g. helfen etc. "to help"
V	e.g. schlafen etc. "to sleep"

These "minus category" features are to be understood as somewhat similar to the slash categories of GPSG, although they are perhaps more akin to the slashes of categorial grammar. The feature indicates that the verb requires a complement of that category in order to become a grammatically complete predicate.

Thus erzählen needs a dative NP and an accusative NP in order to form a grammatically complete predicate.

The device of "minus category" features enables us to replace the large number of VP expansion rules used in previous treatments by one VP expansion rule, shown in (15).

(15) $[_{VP}\ V\ ]$

(15) is not to be understood as a PS rule, but rather as a rule schema that abbreviates a whole series of PS rules that are produced by feature instantiation. The feature instantiation principles applying to (15) produce, among others[5], the instantiated rule (16).

(16)

$[_{VP}$	V	]
[- NP_{dat}]	[- NP_{dat}]	
[- NP_{acc}]	[- NP_{acc}]	

Further, rule schema (17) connects the verb phrases missing complements with their complements.

(17)

$[_{VP}$	X	VP	]
	[- X]		

---

[5] Note that (15) is itself a member of the set of rules produced by feature instantiation from (15).

(17) says that a VP may consist of a category and a VP missing that category. Two instantiations of (17) that are of importance to us are (18 a) and (18 b).

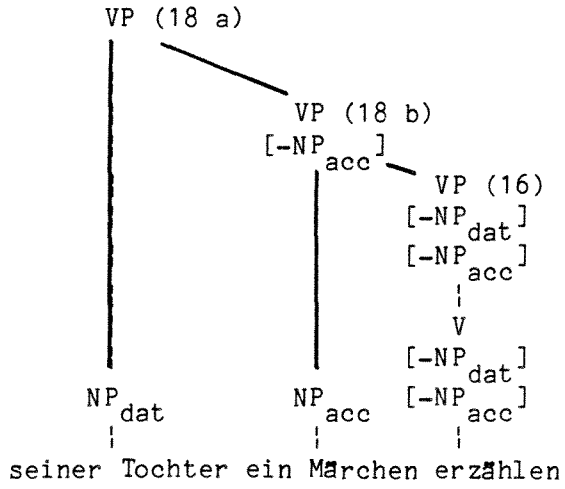
(18 a) $[_{VP}\ NP_{acc}\quad VP\ [-\ NP_{acc}]\ ]$

(18 b) $[_{VP}\ NP_{dat}\quad VP\ [-\ NP_{acc}]\ ]$

$[-\ NP_{acc}]\qquad [-\ NP_{acc}]\ [-\ NP_{dat}]$

(18 a) says that a grammatically complete VP may consist of an accusative NP and a VP missing an accusative NP. (18 b) says that a VP missing an accusative NP may consist of a dative NP and a VP missing both an accusative NP and a dative NP. With these rules we assign the following analysis to the VP in (19), where the number to the right of each node is the number of the rule instantiation that was used to expand each node into its daughters.

(19)



Following Gazdar et al. (1981), I assume that auxiliaries are introduced via a rule schema as shown in (20)[6].

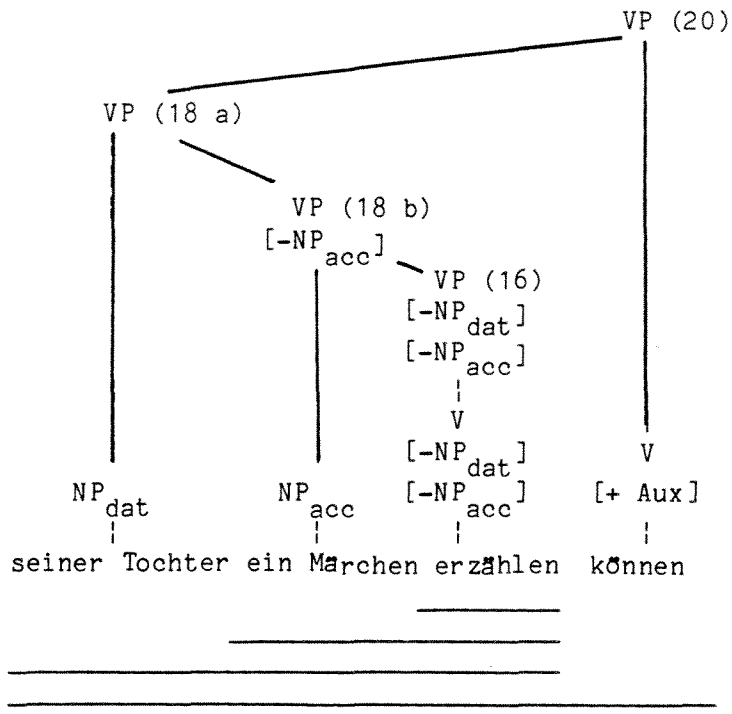
(20) $[_{VP}\ VP\quad V\ [+\ Aux]\ ]$

Using (20), we produce the analysis of a VP consisting of a main verb, auxiliary and direct and indirect objects shown in (21).

---

[6] In Johnson (1982) a precise feature cooccurance system for the German auxiliary is given. The rule schema developed there describes modals, passive and future <u>werden</u> and perfect <u>haben</u> and <u>sein</u>.

(21)



The numbers refer to the number of the rule instantiation employed, and the underlining identifies the strings that are dominated by a VP and hence "frontable" by the rule Fronting[7]. The constituent analysis in (21) corresponds to an analysis according to hypothesis (A) of the last section, and in conjunction with a rule of Fronting would generate (11 a), (11 b) and (11 c).

Assuming that the "minus category" features are foot features in the sense of Gazdar and Pullum (1982), the same feature instantiation principles that gave us rule (16) from (15) give us (22) as an instantiation of (20)[8].

(22) $[_{VP}\ _{VP}\ _{V}\ ]$

$[_{[-NP_{acc}]}\ _{[-NP_{acc}]}\ _{[+\ Aux\ ]}]$

$[_{[-NP_{dat}]}\ _{[-NP_{dat}]}]$

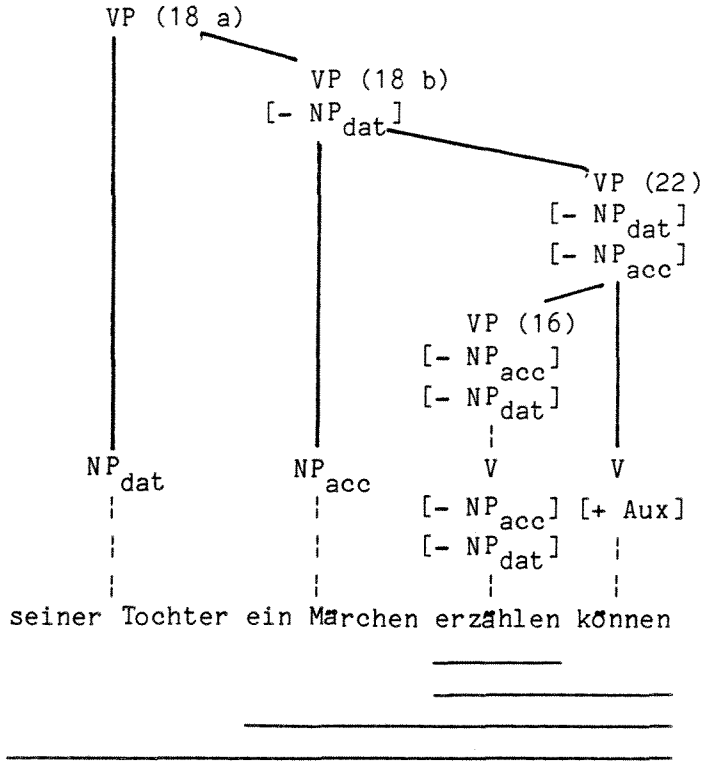
---

[7] The two NPs seiner Tochter and ein Märchen are also "frontable", of course.

[8] The feature instantiations of (20) will also include rules where the "minus category" features trickle down onto the auxiliary. However, by assuming that there exist no lexical elements with features [+ Aux] and [- X], where X is a "minus category" feature, these rules can never be used in a derivation.

With (22), the terminal string in (21) can be reanalysed as in (23).

(23)



The constituent analysis in (23) corresponds to an analysis according to hypothesis (B), and in conjunction with a rule of Fronting would generate (12 a), (13 a) and (13 b).

Similarly, the other feature instantiations of (20) generate the main verb plus object constituents identified by Fronting in (10).

5. Semantics

One objection that might be made to this analysis is that since two alternative derivations are available for the VP in (9), namely (21) and (23), we would expect the clause to be ambiguous, which is not the case. However, as Dowty et al (1981) remark, in system with an explicit semantics, syntactic ambiguity does not imply semantic ambiguity - it is possible that both syntactic analyses are translated into the same semantic representation, and hence are semantically unambiguous. This of course raises the question of how the semantic representation is obtained in the system just described. I describe below two systems for the semantics of "minus category" features which each predict (21) to be synonomous with (23).

The major problem any analysis faces is this: in (21) the auxiliary

kennen maps VP meanings[9] into VP meanings (ie. its semantic type is <VP,VP>) while the auxiliary kennen in (23) maps VPs missings two NPs into VPs missing two NPs. Thus any analysis will have to have a way of resolving this potential "type clash".

One possible solution is to assert that infact the semantic type of VPs with "minus category" features is infact VP, and postulate some other mechanism to link up the object NPs with the verb. We avoid the "type clash" problem by denying that it exists. This approach is taken in the first system presented below, where variable binding is used to connect an object NP with the main verb much in the way variable binding is used to connect the moved element with its "hole" in the Gazdar et al (1981) "slash category" treatment of unbounded dependencies.

A second approach is to assume that VPs with "minus categories" are indeed of different types to a VP without "minus category" features. Other principles must then explain why (21) and (23) are synonomous.

5.1. The Variable Binding Approach

Assuming that auxiliaries as a class are assigned one semantic type, that of mapping VP meanings into VP meanings, ie. <VP,VP>, forces us to postulate that a VP with "minus category" features is semantically a VP. The argument positions in every verb's lexical entry are assumed to be filled with variables of the semantic type corresponding to the syntactic categories the verb subcategorizes for. Thus the lexical entries for the verbs in (14) are as shown in (24).

---

[9] A VP meaning is taken here as an abbreviation for a meaning of type <NP,<s,t>>, ie. for a meaning that needs only one more NP to become a sentence meaning. The semantic type of VPs with "minus category" features is the major topic of the discussion below.

(24)

V [- NP_{dat}] [- NP_{acc}]	e.g.	erzählen, geben etc. "to tell", "to give"
		erzählen'(t_{dat})(t_{acc})
V [- NP_{acc}]	e.g.	lesen etc. "to read"
		lesen'(t_{acc})
V [- NP_{dat}]	e.g.	helfen etc. "to help"
		helfen'(t_{dat})
V	e.g.	schlafen etc. "to sleep"
		schlafen'

In (24), the variables t are of the semantic type NP and the subscripts acc, dat on these variables are indexing mnemonics. The indices are used to ensure that an argument is connected with the correct argument position in the VP[10]. Rule (17), the rule that introduces the "minus categories", can have a semantic translation, given in (25), very similar to that of the metarule introducing slash categories.

(25) $[_{VP}\ X\ \ VP\ ]\ \ \lambda t\ VP'(X')$
$[-X]\ \ [-X]$

where t is of the same type as X, and bears the same index

The semantics of (25) bind the variable inside the VP with the category introduced by the rule. For example, the semantic translations of (18 a) and (18 b) are listed in (26).

---

[10] As Engdahl (1982) notes, some system of indexing of variables is required in a slash category approach to multiple extraction, so the indexing postulated here does not represent an addition to the formal devices postulated.

(26 a) $[_{VP}\ NP_{acc}\ \ VP\ ]\ \ \lambda t_{acc}\ VP'\ (\ NP_{acc}\ )$
$[-\ NP_{acc}]\ \ [-\ NP_{acc}]$

(26 b) $[_{VP}\ NP_{dat}\ \ VP\ ]\ \ \lambda t_{dat}\ VP'\ (\ NP_{dat}\ )$
$[-\ NP_{acc}]\ \ [-\ NP_{acc}]\ \ [-\ NP_{acc}]$
$[-\ NP_{dat}]\ \ [-\ NP_{dat}]$

As mentioned before, the semantic translation of auxiliaries is assumed to be simply a function from VP meanings to VP meanings. (28) shows what the entry for können might look like[11].

(28) V [+ Aux] können' (VP)

We can now procede to derive (21) as in (29). I will show how the meaning of each constituent can be derived from the meaning of its components, but this is merely an expository device, rather than implying I intend the grammar to be "bottom up"[12].

(29)

erzählen -> erzählen'(t_{acc})(t_{dat})

ein Märchen erzählen -> λt_{acc} erzählen'(t_{acc})(t_{dat}) (ein'(Märchen'))
-> erzählen'(ein'(Märchen'))(t_{dat})

seiner Tochter ein Märchen erzählen
-> λt_{dat} erzählen'(ein'(Märchen'))(t_{dat}) (seiner'(Tochter'))
-> erzählen'(ein'(Märchen'))((seiner'(Tochter'))

seiner Tochter ein Märchen erzählen können
-> können'(erzählen'(ein'(Märchen'))(seiner'(Tochter')))

Similarly, we can derive (23) as in (30).

---

[11] I'm interested here in showing how the function argument structure and the lamba mechanism predict synonomy for analyses (21) and (23). Thus I have not given translations for lexical items in terms of Intensional Logic.

[12] I believe that in an account such as this terms such as "top down" and "bottom up" have no empirical content.

(30)
erzählen -> erzählen'(t_{acc})(t_{dat})

erzählen können -> können'(erzählen'(t_{acc})(t_{dat}))

ein Märchen erzählen können ->
λt_{acc} können'(erzählen'(t_{acc})(t_{dat})) (ein'(Märchen'))
-> können'(erzählen'(ein'(Märchen'))(t_{dat}))

seiner Tochter ein Märchen erzählen können
-> λt_{dat} können'(erzählen'(ein'(Märchen'))(t_{dat})) (seiner'(Tochter'))
-> können'(erzählen'(ein'(Märchen'))(seiner'(Tochter')))

Thus the "variable binding" system predicts (21) and (23) to be synonomous, since the same auxiliary is used in both.

The system proposed above does have some drawbacks, in as much as the system of semantic types (see Klein and Sag (1981)) has been largely circumvented by generating verbs with variables in situ, so to speak. This was necessary because a semantic functor (the auxiliary in this case) can only accept one semantic type as its input; thus VPs both with and without "minus categories" must have the same semantic type. Note that the circumvention of the semantic type checking does not result in the production of ungrammatical sentences; the syntactic component of each rule forces a verb to take the correct argument structure[13].

5.2. The Meaning Postulate Approach

The system described above avoided the "type clash" problem mentioned at the beginning of this section by assuming that no type clash existed. In this subsection I assume that the "type clash" exists, and show how one might plausibly propose to deal with it. Assuming the type clash means assuming that the auxiliary können of (21) is of the type <VP,VP>, while the phonetically identical auxiliary of (23) is of type <<NP,<NP,VP>>,<NP,<NP,VP>>>. If we assume that lexical elements only possess one semantic type, and if the grammar of German contains no further specification regarding the nature of the auxiliary, we would regard the auxiliary as a different lexical item to the auxiliary in (23). Then it would be purely fortuitous that (21) and (23) are synonomous, or that the auxiliaries in (21) or (23) simulaneously exist.

Of course it is easy enough to stipulate that the auxilaries in (21) and (23) exist - the appropriate device is the meaning postulate. The meaning postulate in (31) says that for every auxiliary a that

---

[13] There is redundancy between the syntactic rule introducing a verb and its semantic type, which is a lexical property of the verb; since both state the number of objects the verb requires.

maps meanings of type s into meanings of type s, there exists another auxiliary a' that maps meanings of type s missing an NP into meanings of type s minus an NP, and moreover that the two expressions are synonomous.

(31) define the set S of strings[14] recursively as follows
"VP" $\in$ S
if s $\in$ S then "<NP,s>" $\in$ S

let s $\in$ S

let a be a variable over lexical items of type <s,s>
let a' be a variable over lexical items of type <<NP,s>,<NP,s>>

let v be a variable over VP meanings
let n be a variable over NP meanings

Then

$$\forall v \forall n \forall a \exists a' \Box ((a(v (n)) = (a' (v))(n) \text{ and } (\text{phon } (a) = \text{phon } (a')))$$

where phon (x) is the phonological representation of x.

Now of course it is one thing to arbitrarily stipulate something, and something quite different to motivate that stipulation. But consider the case if (31) does not hold - that is, an expression like (23) could be syntactically well-formed, but is ruled out because the auxiliary kennen does not accept arguments of the type <NP,<NP,VP>>. As far as I know, no where else in the grammar are syntactically well-formed sentences ruled out by semantic argument type mismatch, and it seems to be desirable to prevent it from happening here. The argument that an auxiliary takes is already independently specified by the syntax, and it is redundant to specify it again in the semantics. Indeed, a principle of syntactic priority making semantic function-argument structure totally dependent on syntactic structure seems reasonable; ie. if a structure is syntactically well-formed, its semantic argument structure must be well-formed too.

Klein and Sag (1981) give a treatment of control phenomena, which is another area where "type clashes" occur. They noted that these type clashes occured when ever a verb appears in a control structure, and proposed that a Lexical Interpretation Algorithim automatically re-adjusts the controlling verb's semantic argument type to make it fit the available arguments. The LIA as proposed by Klein and Sag accounts for Raising and EQUI structures. Following the principle of syntactic priority, one would extend the LIA to include the type adjustments needed for the German auxiliaries in the class of automatic type adjustments it performs. One way of doing this is to

[14] The quote marks are string delimiters.

incorporate (31) into the LIA; (31) now interpreted as defining the meaning of an auxiliary as a function of its syntactic environment[15].

An alternative to the LIA approach would be to modify the semantic type matching process, perhaps along the lines of Geach (1972), by expanding the type cancellation rules. In standard treatments (eg Dowty et al (1981)) semantic types are reduced according to (32).

(32) For expressions f, g

if type(f) = < a , b >, type(g) = < a >,

then type(f(g)) = < b >

Following Geach, we can supplement (32) with (33).

(33) For expressions f, g

if type(f) = < a , b >,
type(g) = < c_1 ... < c_n , a >...>

then type(f(g)) = < c_1 ... < c_n , b >...>

where f(g) $=_{def}$ λ x_1 ... x_n f(g(x_1 ... x_n))

and x_i, 0 < i < n+1, is a variable,
and type(x_i) = < c_i >

With this extension to the type theory, the type clashes disappear. For example, in (23) the modal können of type <VP,VP> applies to the verb erzählen of type <NP,<NP,VP>>. By (33), with f = können and g = erzählen, we predict the semantic type of the constituent erzählen können to be well formed with semantic type <NP,<NP,VP>>, and, since the object NPs are lambda bound to the main verb by virtue of the second part of (33), we assign (21) and (23) identical translations.

Thus assuming that the syntax completely determines the semantic function-argument structure, we predict (21) and (23) to be synonomous.

6. Conclusion

At the beginning of this paper I presented data which suggested that any grammar of German would need to provide at least two analyses of German VPs consisting of objects, main verb and auxiliary. But given the assumption that the representation of verbs in the lexicon is of the form in (14), the principles of feature instantiation predict

---

[15] That is, a and a' are interpreted as different semantic functors associated with a single lexical item, rather than different lexical items.

that VPs incorporating both auxiliaries and objects will have multiple analyses unless further stipulations are added to the rule schema that introduces auxiliaries. It is strong confirmation of these principles that they predict exactly the correct constituent analysis with respect to fronting.

Of course the analysis presented here is just the beginning. The interaction between the "minus category" and the slash category features needs to be further investigated: in particular, is it just a coincidence that both are foot features and both can be described using the similar systems of semantic interpretation, or is there a generalization here that we have missed?

Secondly, what is the status of the principle that the function argument structure is strictly determined by the syntactic environment? If the second proposal for the semantic analysis of minus categories is adopted, the structure of the Lexical Interpretation Algorithim is 3 disjunctive statements. It would be nice if the separate statements inside the LIA could be unified.

Finally there is the question of the utility of the "minus categories". It seems that they provide a natural way to generate multiple syntactic analyses. It would be interesting to see if phenomena described in other frameworks as due to "re-analysis" can be described with them.

7. Bibliograophy

Bierwisch M (1965) Grammatik des deutschen Verbs; Studia Grammatika II; Akademie Verlag, Berlin

Dowty D, Wall R and Peters S (1981) Introduction to Montague Semantics; Reidel, Dordrecht.

Engdahl E (1982) Constituent Questions, Topicalization and Surface Structure Interpretation in GPSG in Proceedings of the first West Coast Conference on Formal Linguistics, Stanford Linguistics Association.

Gazdar G, Pullum G and Sag I (1981) Auxiliaries and related phenomena in a restrictive theory of grammar, Indiana University Linguistics Club, Indiana

Gazdar G and Pullum G (1982) Generalized Phrase Structure Grammar; A Theoretical Synopsis, Indiana University Linguistics Club, Indiana

Geach P T (1972) A program for syntax in Semantics of Natural Language; Reidel, Dordrecht

Johnson M (1982) The structure of the German VP unpublished ms. UCSD

Klein E and Sag I (1981) Semantic Types and Control, presented at LSA, New York

Kunsman P (1973) Verbale Gefüge, Linguistische Reihe, Band 14, Max Hueber Verlag, München

Nerbonne J (1982) "Phantoms" and German Fronting: Poltergeist Constituents?, presented at LSA San Diego

Uszkoreit H (1982) German Word Order in GPSG in Proceedings of the first West Coast Conference on Formal Linguistics, Stanford Linguistics Association.