German partial VP fronting[1]
Richard Hudson

1.   Overview of the problem and the solution
The problem lies in German sentences like the following:
(1)   a.   [Er] darf jeden Tag Kaffee trinken.
            He may every day coffee drink
            `He can drink coffee every day.'
        b.   [Kaffee] darf er jeden Tag trinken.
        c.   [Jeden Tag Kaffee trinken] darf er.
        d.   [Jeden Tag trinken] darf er Kaffee.
        e.   [Eine Concorde gelandet] ist hier noch nie.
            A Concorde landed is here still never
            `A Concorde has never yet landed here.'
        f.   [Ein wirklicher Fehler unterlaufen] war ihm noch nie.
            a real mistake happened was to.him still never
            `He had never yet made a real mistake.'
        g.   *[Er trinken] darf jeden Tag Kaffee.
The sentences illustrate the V2 character of German which allows precisely one
constituent before the finite verb. We can call this first constituent simply `X', and
enclose it in square brackets.

        In the first three sentences X is straightforward - the subject, the object or
the whole VP `drink coffee'. Example (d) is more interesting because X is just the
verb `drink' plus an adjunct, `every day' - i.e. less than the maximal projection
expected in the topic position; and more interestingly still, (e) and (f) show that X
can even include the fronted verb's subject but not its complements and adjuncts -
though the ungrammaticality of (g) alerts us to the fact that this pattern is
constrained in some way.

        These patterns have attracted a great deal of interest in the past decade,
partly because of the interesting challenge they present for any syntactic theory,
and partly because the data seem to be rather clear and uncontroversial even
down to fine details, which allows discussion to focus on questions of theory and
interpretation. Section 2 will survey earlier work, but the main characteristic of this
earlier work is easily stated: it has all been done in terms of theories whose
foundation is phrase structure. According to all these analyses, X in the interesting
examples is a `partial VP', so the construction is called `Partial-VP Fronting'
(**PVPF**). We shall see that the debate has produced a range of interesting solutions
which each solve most of the problems, but they all have weaknesses which (I
shall suggest) are inherent to the phrase-structure approach.

        Suppose we abandon phrase structure (**PS**) as the basis for grammar.
Instead of building a sentence out of complex building blocks like NP and VP, we
could take a much more traditional approach in which the only building blocks are
single words. This is the essence of **dependency structure** (**DS**), in which
sentence structure is defined in terms of dependencies between individual words;
this approach underlies a long and rich European tradition of grammatical analysis,
though it has had much less influence in North America. Early in the recent more
theoretical tradition (e.g. Tesnière 1959/1965) dependency analysis focussed on
semantic dependencies without paying much attention to word order, so PVPF is

no problem at all: it is easy to draw a dependency diagram which links up all the words in a meaningful way. But this is a theoretical disaster, because PVPF is unusual and ought to challenge any theory of sentence strucure. After all, English does not allow PVPF at all, as witness the examples in (2).

(2)   a.   Drink coffee he may, but he mustn't drink chocolate.
       b.   *Drink he may coffee, but he mustn't make it.

Why is PVPF allowed in German but not in English? And why can German speakers co-front the subject in some cases but not in others? A blanket acceptance of any word order does not explain anything, so traditional dependency analysis complements phrase-strucure analysis nicely: whereas PS analysis finds PVPF too hard, traditional DS analysis finds it too easy. Nevertheless, the DS approach offers an interesting alternative to PS precisely because the problems change.

The literature which I shall review has been formulated in terms of one of two theoretical paradigms: the Chomskian tradition which has moved from Government-and-Binding theory (Chomsky 1981) to the Minimalist Program (Chomsky 1995; henceforth **GB/MP**) and Head-driven Phrase Structure Grammar (**HPSG**; Pollard and Sag 1994). The WG analysis will combine leading ideas from both traditions. In a nutshell, I shall follow HPSG in treating the non-fronted dependents of the lower verb as `structure-shared' with the higher verb. For example, in Trinken darf er Kaffee, `Drink he may coffee', I shall suggest that `coffee' depends on `may' as well as on `drink' and a general principle (the Raising Principle) will require it to take its position from `may' rather than from `drink'; in contrast, in Kaffee trinken darf er, `coffee' only depends on `drink' so the position of `may' is irrelevant. English does not allow objects to be structure-shared in this way, which is why we do not allow *Drink he can coffee.

As with other theories, the big problem for a WG analysis is the co-fronting of subjects. In this case the start of my solution will be similar to the one favoured by GB/MP analysts, namely to give subjects the same status as other dependents. Whereas GB/MP analyses achieve this by including the subject in the VP, my WG analysis will simply generalize across all dependents, including subjects. But this raises a second problem: why does the subject of a supposedly `raising' verb take its position from the lower verb? In other words, why does it behave as though it was lowered? Why doesn't the Raising Principle mentioned above prevent this? The answer will be that it is in fact lowered, but that this can be treated as an explicitly permitted exception to the Raising Principle. Co-fronted subjects also raise a second set of problems: why do some subjects resist co-fronting? In this case the WG analysis will involve the interface between syntax and semantics, and will explain the restrictions in terms of the impossibility of giving the fronted verb a stable meaning.

The approach that I shall take is based on **Word Grammar (WG** - see Hudson 1984, 1990, 1993, 1994a, 1995a, b, 1997a, b), which uses an enriched version of DS for syntactic analysis in combination with **default inheritance**[2]. DS allows rules and principles to apply directly to individual word-word dependencies rather than to whole phrases; and default inheritance allows not only words but also dependencies to be exceptions to the `default' generalizations. Various constraints on word order make it almost (but not quite) equivalent to PS, and some of the ideas found in the various PS solutions to PVPF will translate more or

less directly into WG. However there is enough difference between WG and PS to allow a different range of solutions, which (I shall argue) bring us nearer to a complete understanding of PVPF.

These ideas may be easier to grasp, even at this stage, in the diagrams of **Figure 1** which I shall explain later.
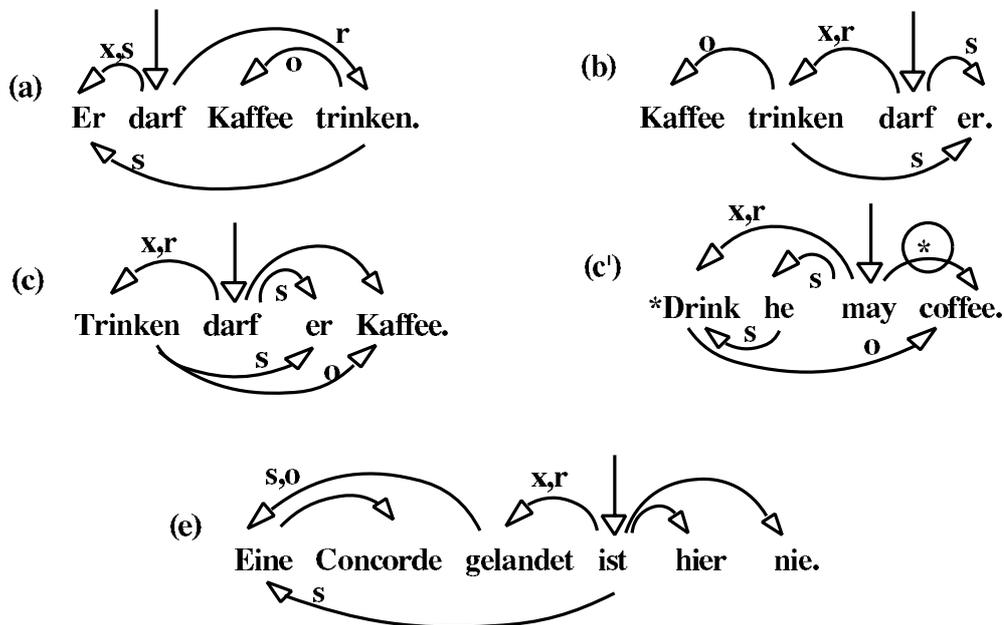

Figure 1

2.   A pre-theoretical survey of VP-fronting

This survey will move from general to particular - from fronting in general to fronting of `verb phrases', and then more specifically to verb phrases which are fronted together with their subjects. Most of the facts are more or less well-known, but I shall anticipate the later discussion by expressing them in terms of dependency structure. The generalizations that we shall try to model more formally are summarised in (25) at the end of this section.

By `fronting' I mean the choice of a main clause's `X' - its first constituent, traditionally called its `Vorfeld' (`fore-field'). As a strict V2 language, German requires any declarative non-dependent finite verb to have precisely one preceding dependent, its `X' (which we might also call its topic). The possibilities are illustrated by the following examples from Uszkoreit (1987:414), where the finite verb is the modal verb <u>sollte</u>, `was to' and X is bracketed.

(3)   a.   [Der Kurier]$_X$ sollte den Brief nachher einem Spion zustecken.
              the courier should the letter afterwards to.a spy slip
              `The courier was to slip the letter afterwards to a spy.'
       b.   [Den Brief]$_X$ sollte der Kurier nachher einem Spion zustecken.
       c.   [Einem Spion]$_X$ sollte der Kurier nachher den Brief zustecken.
       d.   [Nachher]$_X$ sollte der Kurier den Brief einem Spion zustecken.
       e.   *[Der Kurier den Brief]$_X$ sollte nachher einem Spion zustecken.
       f.   *[Den Brief nachher]$_X$ sollte der Kurier einem Spion zustecken.

What concerns us in this paper is the possibility of fronting the non-finite verb (<u>zustecken</u>, `to slip') without all of its dependents: PVPF.

The freedom of word order in PVPF is quite striking to an English speaker, and is clearly linked to the freedom of word order in full clauses (so-called `scrambling', to which we shall return in the next section); according to den Besten and Webelhuth (1990:77) German and Dutch are the only Germanic languages that allow PVPF, and also the only ones that allow scrambling. Uszkoreit (1987) illustrates the freedom of PVPF with further permutations on the basic sentence given above:

(4)   a.   [Zustecken]$_X$ sollte der Kurier den Brief nachher einem Spion.
         b.   [Den Brief zustecken]$_X$ sollte der Kurier nachher einem Spion.
         c.   [Einem Spion zustecken]$_X$ sollte der Kurier nachher den Brief.
         d.   [Nachher einem Spion zustecken]$_X$ sollte der Kurier den Brief.

It seems to be generally agreed[3] that virtually any imaginable combination is grammatical, with just two clear exceptions. In particular, there is no difference between complements and adjuncts. In dependency terms (with these exceptions) any dependent of a fronted non-finite verb may (but need not) be `co-fronted' with it. This very simple statement illustrates one of the attractions of dependency analysis: the possibility of generalizing across all dependents without distinguishing complements from adjuncts.

What is surprising about these examples is not the possibility of co-fronting dependents; after all, this is exactly what one finds in English to the extent that non-finite verbs can be fronted. Rather, the surprise lies in the possibility of separating a fronted verb from some of its dependents, including its complements. The choice of not co-fronting complements simply does not exist in English:

(5)   a.   .. so slip a letter to a spy he did.
         b.   *.. so slip he did a letter to a spy.
         c.   *.. so slip a letter he did to a spy.
         d.   *.. so slip to a spy he did a letter.

The freedom is actually even greater than is implied in the above discussion, because it is not just dependents of the fronted verb that may be separated from it, but even dependents of its dependents. Admittedly this recursion is limited to further non-finite dependents, but it applies quite generally to them: the dependents of a `coherent[4]' non-finite verb which is itself co-fronted with a higher non-finite verb may follow the finite verb. The following examples are based on Hinrichs and Nakazawa 1994, Müller 1995 and Prinzhorn 1990:

(6)   a.   [Peter]$_X$ wird Maria das Buch geben können.
                 Peter will to.Mary the book give can
                 `Peter will be able to give Mary the book.'
         b.   [Maria das Buch geben können]$_X$ wird Peter.
         c.   [Geben können]$_X$ wird Peter Maria das Buch.
(7)   a.   [Er]$_X$ wird seiner Tochter ein Märchen erzählen müssen.
                 he will to.his daughter a fairy-tale tell must
                 `He will have to tell his daughter a fairy-tale.'
         b.   [Seiner Tochter ein Märchen erzählen müssen]$_X$ wird er.
         c.   [Erzählen müssen]$_X$ wird er seiner Tochter ein Märchen.
(8)   a.   [Zu verletzen versucht] hat Peter den Paul niemals.

to injure tried has Peter the Paul never
        `Peter has never tried to injure Paul.'
    b.   *[Zu verletzen bedauert] hat Peter den Paul niemals.
        to injure regretted has Peter the Paul never
In example (6), both `Mary' and `the book' depend on `give', and as we might
expect, if `give' is co-fronted with `can' they can co-front too; but they also have the
option of not co-fronting, as though they depended on the finite verb itself. Similar
remarks apply to the examples in (7) and (8), except that the latter shows that the
lower verb need not be a bare infinitive, but must be coherent (as with `try' but not
with `regret').

        As I mentioned earlier, there are two special restrictions that have to be
placed on this general freedom either to co-front or to follow the finite verb. The
first restriction applies to the fronted verb pairs in the last examples. In these
examples, one non-finite verb is the coherent complement of another, and the
second is fronted; for example `give' is the complement of `can', which is fronted.
In such cases it seems that the verb which is complement of the fronted verb does
not have the usual option of following the finite verb; it must be co-fronted. Hinrichs
and Nakazawa give the following judgement on one permutation of example (7);
the other examples are based on Baker 1994 and Uszkoreit (personal
communication).

(9)   *[Müssen]$_X$ wird er seiner Tochter ein Märchen erzählen.
(10)   a.   [Das Examen bestehen]$_X$ wird er können.
                the exam pass will he can
                `He will be able to pass the exam.'
        b.   [Das Examen bestehen können]$_X$ wird er.
        c.   *[Können]$_X$ wird er das Examen bestehen.
(11)   a.   [Den Apfel gegessen]$_X$ wird er haben.
                the apple eaten will he have
                `He will have eaten the apple.'
        b.   [Den Apfel gegessen haben]$_X$ wird er.
        c.   *[Haben]$_X$ wird er den Apfel gegessen.

One generalization that would cover all these examples is that a non-finite
complement of a non-finite verb must be immediately before it; in Hinrichs and
Nakazawa's terms, a non-finite verb and its non-finite complement must
(sometimes) form an inseparable `verbal complex': bestehen können, gegessen
haben.

        These verbal complexes are important because they play a part in another
area of grammar, the ordering of clause-final verbs. Subordinate verbs generally
follow their dependents; we have already seen this pattern in PVPF, where X
follows any co-fronted dependents (e.g. `the apple eaten'). If a non-finite verb
depends on a verb which is itself dependent, the two generally stack up at the end
of their clauses with the lower verb (as expected) before the higher one:

(12)   Ich wusste, daß er das Examen bestehen können würde.
        I knew that he the exam pass can would
        `I knew that he would be able to pass the exam.'

There is an exception, however: what is called `modal flip' (Baker 1994), `auxiliary
flip' (Hinrichs and Nakazawa 1994) or `infinitive as past participle' (Haider et al
1995:11). This allows a verbal complex to follow the finite auxiliary on which it

depends; so alongside the order in (12) we also find the following:

(13) a.   Ich wusste, daß er das Examen würde bestehen können.
     b.   *Ich wusste, daß er das Examen bestehen würde können.
     c.   *Ich wusste, daß er das Examen würde können bestehen.

These facts follow from our earlier generalization about verbal clusters: if one non-finite verb is the complement of another, the first must be immediately before the second. This rule deals with the first of our two exceptional restrictions on the general freedom of X's dependents to either co-front or not.

The second restriction concerns X's subject. In English it is impossible to co-front a subject, and the same is true for many German sentences:

(14) a   *.. so [he pass a letter] did.
     b   *[Der Kurier zustecken]$_X$ sollte den Brief nachher einem Spion.

This is exactly as predicted by a VP analysis in which subjects are outside the VP, though as we shall see this apparent benefit of PS has been lost in some recent analyses. Dependency analysis, on the other hand, does not - and cannot - recognise a VP, so the ban on subject fronting calls for an explanation. I shall introduce a general principle which will achieve this effect, but both explanations face the same fundamental problem: some subjects in German can be co-fronted with the non-finite verb. This rather surprising fact is illustrated in the following examples which Uszkoreit (1987:415) attributes to Haider (and which according to Haider (1993:151) are already anticipated in Haider 1981).

(15) a.   [Ein wirklicher Fehler unterlaufen]$_X$ war ihm noch nie.
            a real mistake happened was to.him still never
            `He had never yet made a real mistake.'
     b.   [Solch eine Frau begegnen]$_X$ kann auch nur dir.
            such a woman meet can also only to.you
            `Only you can run into such a woman.'

In example (a) `a real mistake' is certainly the subject of `happened' - this is clear from its nominative case and from the verb agreement in a simple sentence:

(16) a.   Ein Fehler unterlief ihm.
            `He made a mistake.'
     b.   Zwei Fehler unterliefen ihm.
            `He made two mistakes.'

Nor is there any doubt that `a real mistake' is the subject of `was', which agrees with it[5]:

(17) Zwei Fehler unterlaufen waren ihm noch nie.
            `He had never made two mistakes.'

Similar remarks apply to `such a woman' in example (b). In both cases, therefore, the subject which is co-fronted with the non-finite verb is also the subject of the finite verb.

Why is subject co-fronting allowed in these two examples but not in (14b)? Because they fall on opposite sides of a distinction that the grammar draws between subjects that are co-frontable and those that are not. The following survey of constraints is based primarily on Haider (1990). First, to be co-frontable a subject must be indefinite; the following are Haider's examples:

(18) a.   Ein/??Dieser Fehler unterlaufen ist ihr noch nie.
            a/this mistake happened is to.her still never
            `She has never yet made a/??this mistake.'

b.   Ein/??Der Außenseiter gewonnen hat hier noch nie.
       a/the outsider won has here still never.
       `An/??the outsider has never yet won here.'

The second set of constraints involve the fronted verb's transitivity. Roughly speaking, it should have low transitivity (in the sense of Hopper and Thompson 1980), so the co-fronted subject should be low on the `activity' scale, and the object, if distinct from the subject, should be relatively unimportant. This generalization is intended to cover the following three more specific cases that have been vouched for by native speakers.

■ The fronted verb is unaccusative - i.e. in German a verb that selects the auxiliary <u>sein</u>, `be', such as <u>unterlaufen</u>, `happen' (of mistakes), <u>entweichen</u>, `escape' or <u>landen</u>, `land' as in example (1e) repeated here:

(19)   [Eine Concorde gelandet]$_X$ ist hier noch nie.
        A Concorde landed is here still never
        `A Concorde has never yet landed here.'

Unaccusative verbs may have co-fronted subjects even if they also have arguments that follow the finite verb:

(20)   a.   [Ein Fehler unterlaufen]$_X$ ist auch schon mal diesem Professor.
              a mistake happened is also already once to.this professor
              `This professor did in fact once make a mistake.'
       b.   [Ein Tiger entwichen]$_X$ ist doch erst kürzlich diesem Wanderzirkus.
              a tiger escaped is though just recently to.this travelling.circus
              `A tiger has just recently escaped from this travelling circus though.'

■ The fronted verb is passive. It appears that passive verbs also allow co-fronted subjects under the same conditions as unaccusatives - which is to be expected according to standard analyses if both passive and unaccusative subjects are also objects:

(21)   a.   (Und rate mal, was dann gemacht wurde.) [Ein Witz erzählt]$_X$ wurde[6].
              and guess just what then done was - a joke told was
              `And just guess what happened then. A joke was told.'
       b.   [Ein solch schönes Geschenk gemacht]$_X$ wurde mir noch nie.
              a such beautiful present made was to.me still never
              `Such a beautiful present was never yet given to me.'

■ The fronted verb is `complete'. This is my name for a third category that Haider recognises: a verb that has no complement (`argument' in Haider's terms) other than a pronoun[7]. Rather surprisingly this allows even transitive verbs to have co-fronted subjects, provided only that their object is a pronoun immediately after the finite verb. The following are Haider's examples:

(22)   a.   [Ein Außenseiter gewonnen]$_X$ hat (das) da noch nie.
              an outsider won has (that) there still never
              `An outsider has never yet won (that) there.'
       b.   [Linguisten gespeist]$_X$ haben (das) dort noch nie.
              linguists eaten have (that) there still never
              `Linguists have never yet eaten (that) there.'
       c.   [Kinder gespielt]$_X$ haben (das) hier noch nie.
              children played have (that) here still never
              `Children have never yet played (that) here.'

Similarly, Webelhuth (1990:53) describes (23) as `more or less fully grammatical'.

7

(23)   [Leute getanzt]$_X$ haben hier noch nie.
        people dance have here still never
        `People have never yet danced here.'

In contrast, a full-blown object is ungrammatical:

(24)   a.   *[Ein Außenseiter gewonnen]$_X$ hat da noch nie das Derby.
               an outsider won has there still never the Derby.
               `An outsider has never yet won the Derby there.'

        b.   *[Linguisten gespeist]$_X$ haben dort noch nie Langusten.
               linguists eaten have there still never lobsters.
               `Linguists have never yet eaten lobsters there.'

        c.   *[Kinder gespielt]$_X$ haben hier noch nie Tempelhüpfen.
               children played have here still never (?temple hops?)
               `Children have never yet played (?temple hops?) here.'

How can we make sense of this curious detail? Why should a pronoun attached to the finite verb be permitted whereas a full NP is not? I have to admit in advance that I have no satisfactory answer, but two facts seem relevant. One is that this is the normal position for a pronoun even if it belongs to a subordinate verb[8]; so `that' after `has' in (22a) `counts' as though it were next to `won'. The second fact is that in `won has that' the pronoun is physically very close to `won', which allows it to be processed much sooner than `the Derby' in (24a).

      This completes the pre-theoretical survey of the facts relating to PVPF. They are summarised below:

(25)   **Partial VP-fronting: the facts**
        a. A declarative non-dependent finite verb must have just one preceding dependent, its `X'.
        b. If the X is a non-finite verb, any of its dependents may either be co-fronted with it or follow the finite verb.
        c. If another non-finite verb depends on a non-finite X, its dependents may either be co-fronted or follow the finite verb just like those of X itself.
        d. If one non-finite verb (V1) is the complement of another (V2), V2 must immediately follow V1.
        e. The subject of a non-finite X (which is also the subject of the finite verb) may be co-fronted provided that:
        ▪ it is indefinite, and
        ▪ it has low transitivity - unaccusative, passive or free of non-pronoun complements.

3.   Phrase-structure analyses of PVPF

In the last decade PVPF has been the focus of a good deal of attention within two communities of linguists, both of whom assume some version of phrase structure (rather than dependency structure). One community is part of the Chomskian tradition which has moved from Government-and-Binding theory (Chomsky 1981) to Minimalism (Chomsky 1995), while the other is in the monostratal tradition which started in Generalized Phrase Structure Grammar (Gazdar et al 1985) and now uses Head-driven Phrase Structure Grammar (Pollard and Sag 1994). These two communities have been discussing very similar issues, but rather surprisingly there is little evidence in the literature that they are even aware of each other's existence.

      Both communities start from the same problem: PVPF requires the grammar

to be able to recognise syntactic units which at least appear not to be constituents which would otherwise be recognised. The facts point clearly to some kind of link between PVPF and the free order of elements in full clauses, so both communities have looked for a solution in which the non-standard units of PVPF are generated as a by-product of the mechanisms that allow free word order. In both communities various alternative solutions have been suggested and the debate is still in progress. It would be presumptuous for an outsider to offer an evaluation of the alternatives when the community itself cannot agree, so I shall simply outline the solutions and comment on some ways in which the phrase-structure assumptions seem to contribute to the problem rather than to its solution.

In the Chomskian community there are two main views. The majority view seems to favour a combination of `scrambling' and fronting (i.e. movement to Spec of C). The main exponents of this view are den Besten and Webelhuth (1990) and Webelhuth (1990), but roughly the same analysis is supported by Grewendorf (1988). The leading idea is that scrambling produces whatever VP's are needed for fronting, including those that are `partial'. The following examples are based on a pair from den Besten and Webelhuth (1990:78):

(26) a   weil Hans [$_{I'}$ das Buch$_i$ [$_{I'}$ nicht [$_{VP}$ t$_i$ gelesen] hat]]
              because Hans the book not read has
              `because Hans has not read the book'

      b   [$_{VP}$ t$_i$ gelesen]$_k$ hat Hans [$_{I'}$ das Buch$_i$ [$_{I'}$ nicht vp$_k$]]
              `Hans has not read the book.'

Example (b) is grammatical because the fronted VP is already available in example (a), thanks to scrambling. This analysis has the attraction of explaining the link between scrambling and PVPF which we noted earlier, but it stands or falls with scrambling.

From a dependency viewpoint, scrambling is a rule whose main role is to undo the effects of phrase structure. Phrase structure, particularly in the Chomskian tradition, makes fixed ordering of elements normal so extra machinery must be invoked to handle free order; whereas the reverse is true of dependency structure. In the absence of any extra constraints, the dependents of a word may occur in any order. This theoretical difference does not in itself mean that dependency structure is better than phrase structure for free-order constructions; maybe it will turn out that we need the extra structure provided by the phrases that scrambling builds (e.g. in order to explain subject-object asymmetries). Nor does dependency structure in itself immediately solve the PVPF problem; as we shall see in section 6 the problem is almost the same for both approaches. However the fact that scrambling is needed at all raises serious questions about the fundamental role that phrase structure is normally given.

There is another reason for being worried about scrambling. There is great uncertainty even among its advocates about what kind of phenomenon scrambling is. As Haider et al (1995:15) show, almost all the logical possibilities have been considered, but none have been decisively rejected. Are scrambled structures base-generated or created by a movement transformation? And in the latter case, does it adjoin or substitute, and is the `landing site' an A or an A-bar position? Which categories may be scrambled, and which may be landing sites? Apart from these descriptive uncertainties there are more general theoretical concerns. How does this transformation fit into a theory which nowadays eschews parochial

transformations? Worse still, how does it mesh with the Minimalist assumption that all movement is forced by morphological features[9]? With so much uncertainty about scrambling we cannot be sure that the details of scrambling can ever be sorted out satisfactorily; indeed, we can't even be sure that scrambling, as such, is a real phenomenon.

We should also recognise a minority view in the Chomskian tradition which is nevertheless influential, that of Haider (1990, 1993), according to whom fronted VPs are base-generated as full VPs in the Specifier position of Comp. The fronted VP is integrated with the rest of the sentence by means of indices which link it to a slot for a single verb - a V^0 slot within the `verb complex' (VC). The same analysis covers co-fronted subjects, so subjects are also assumed to be part of VP. Haider is inexplicit about the link between the fronted VP and the parts that follow the finite verb, but one assumes that these are somehow co-indexed with traces in the fronted VP. The following examples are based rather loosely on the discussion in Haider (1990:104, 109), where partial structures are offered:

(27)   a   [Mir geholfen]$_i$ hat er nicht [$_{VC}$ [e$_i$] e$_j$]]
              me helped has he not
              `He has not helped me.'

       b   [[e$_k$ e$_j$ Geöffnet]]$_m$ hat$_i$ sie ihm$_k$ die Augen$_j$ [$_{VC}$ [e$_m$] e$_i$]].
              opened has she to.him the eyes
              `She has opened his eyes.'

If the structure in (b) is a correct interpretation of Haider's intentions, the traces inside the fronted VP must be sanctioned by the same principles as apply in scrambling, so the general worries about scrambling apply here too.

The other community working on PVPF have used some version of Head-driven Phrase Structure Grammar (**HPSG**: Pollard and Sag 1994). The HPSG tradition starts with three preliminary discussions in terms of Generalized Phrase Structure Grammar (Nerbonne 1986, Johnson 1986 and Uszkoreit 1987) and includes the following: Netter 1991, Baker 1994, Hinrichs and Nakazawa 1994, Nerbonne 1994 and Müller 1995.

Netter introduces the general idea that runs through all the other analyses, although he does not try to formalise it: the idea that PVPF shows some kind of `clause union' in which the two verbs share their dependents. Baker develops this idea by suggesting (1994:33) that the two verbs should structure-share all their complements, just as they do their subjects. This gives a completely flat structure for most clauses containing multiple verbs such as the following example:

(28)   .. daß du uns die Schlacht gewinnen helfen wirst
              that you us the battle win help will
              `that you will help us to win the battle'

The mechanism for showing this is the HPSG equivalent of scrambling, but it is crucially different because it involves operations on the `COMPS' lists for the various verbs whereby the dependents of the lower verbs also appear among the complements of the higher ones.

              complements of `win': `the battle'
              complements of `help': `us', `the battle', `win'
              complements of `will': `you', `us', `the battle', `win', `help'

These dependencies parallel those in which `you' is structure-shared as subject by all three verbs; this subject-to-subject raising is found in English as well, but

Baker's analysis of German generalizes it to all dependents so we can call it
`generalized raising[10]'.

Generalized raising explains why some dependents can stay after the finite
verb if a lower non-finite verb is fronted. If a phrase is a complement of both V1
and V2, and V2 is higher than V1, `raising' requires the phrase to take its position
among the dependents of V2 rather than of V1. In a curious reversal of the normal
problem this analysis leaves us without an explanation for the more English-like
option of co-fronting, in which the shared phrase takes its positions from V1. To
allow co-fronting Baker suggests a special rule for PVPF (1994:38). Baker
illustrates the analysis with example (29).

(29)   [Füttern dürfen]$_X$ wird Cecilia das Nilpferd.
        feed may will Cecilia that hippo
        `Cecilia will be allowed to feed the hippo.'

X is recognised as a constituent belonging to the special category `PVP', and the
dependencies are shown, as above, in the SUBJ and COMPS lists of the various
verbs:

        complements of `feed': `the hippo'
        complements of `may': `feed', `the hippo'
        complements of `will': `Cecilia', `the hippo'

(The analysis of `will' does not translate easily into dependency terms.)

Baker's analysis captures a basic insight which I shall develop in my
dependency-based analysis: the complex word-order facts reflect complexities in
the dependency relations. When a dependent of the fronted verb does not co-front
with it, this is because it is linked to the finite verb by a dependency which is
similar to the dependency that links a raised subject to a higher verb. This
dependency has nothing to do with meaning - the hippo is not a semantic
argument of `may' or `will', for example, just as Cecilia is semantically not related to
`will' although the word Cecilia is clearly the subject of the word wird (as witness,
among other things, the agreement pattern that links them). This idea is applied, in
one form or another, in all the HPSG analyses of PVPF.

Baker also addresses several of the specifics of PVPF. Her main concern is
`modal flip', which she handles through subcategorisation details for the auxiliaries
concerned. She also discusses co-fronted subjects, which she suggests may be
generated if the subjects concerned are treated as complements. This analysis has
some attractions in explaining co-fronting of unaccusative subjects though it does
not generalise to Haider's transitive and non-ergative examples.

Alongside these strengths, though, we must recognise that Baker's analysis
does not work for adjuncts. This is a serious weakness which it shares with other
HPSG analyses. Raising and co-fronting are both handled by operations on the
`COMPS' list, but (by definition) this does not contain adjuncts. But we have seen
that complements and adjuncts are treated alike by PVPF, so we really need a
single analysis which applies equally to both. This weakness is shared by Hinrichs
and Nakazawa 1994, as Müller points out (1995), but Müller does not explain how
his own analysis solves the problem. Any analysis which rests on the manipulation
of lexical valencies is incapable of generalizing beyond these to adjuncts. In a
dependency-based analysis, on the contrary, it is very easy to generalize across
adjuncts, complements and subjects: they are all simply `dependents'.

To summarise the achievements of these two communities, it would be fair

to say, I think, that all their successes have required a shift from standard PS assumptions in the direction of DS. GB/MP analyses build on scrambling, which is problematic in PS but trivial in DS; and they explain co-fronted subjects by giving subjects and objects a similar status, just as in DS. On the other hand, HPSG analyses invoke `structure sharing', which means sharing of dependents, rather than scrambling; but at present this only applies to complements. The analysis needs to be extended in two directions: to cover subjects (when co-fronted) and adjuncts when not co-fronted. This will involve embracing the DS concept of `dependent', which cannot be captured in current HPSG theory. The next section will introduce a theory of grammar in which the role of phrase structure is nill (or more precisely, greatly restricted) and dependencies do (almost) all the work. The remaining sections will apply this theory to most of the facts about PVPF.

4.   Enriched dependency structure and raising
The alternative to phrase structure is dependency structure, in which the fundamental syntactic relationship is not the part-whole relationship of a word (or phrase) to its mother phrase, but the whole-whole dependency relationship between two words in which one word (the dependent) depends on another (its `**parent**[11]'). The two systems translate roughly into each other (though there are enough differences to prevent them from being simply notational variants). PS implies DS (to the extent that phrases are endocentric), and DS implies PS because each word defines a phrase consisting of it (as head) and the phrase of each of its dependents. If A is the head of a phrase which is a non-head (complement, adjunct or specifier) in a phrase whose head is B, then A depends on B; and vice versa.

        One of the differences between PS and DS is in the treatment of discontinuous phrases. In PS continuity of phrases is built into the theory because phrases are (by definition) equivalent to bracketings of the string of words. A discontinuous phrase cannot be shown by bracketing. In DS, on the contrary, continuity has to be imposed by general principle as an optional extra. In the DS tradition there is a principle of `projectivity' which enforces continuity, and which is accepted by some theorists and rejected by others (Fraser 1994). It seems clear that phrases normally are continuous and that most discontinuous phrases are ungrammatical (compare with very great difficulty and *very with great difficulty), so there should be a general constraint which bans it; but there are some discontinuous phrases that are grammatical, so given suitable constraints the relative flexibility of DS allows just the right amount of flexibility. For example, subject-to-subject raising may involve a discontinuous phrase consisting of the raising verb's subject plus its complement such as It ... raining in It was raining. Here discontinuity is allowed, so DS needs a mechanism which will allow discontinuity in some cases while forbidding it in general. What follows is an account of how this is done in one particular manifestation of DS theory, Word Grammar.

        Unlike most other DS theories, WG allows a word to depend on more than one other word. In It was raining, for example, it depends on both the verbs. However, the dependencies concerned carry different types of relationship. The link to was is relevant to word order and to verb agreement, while the one to raining is responsible for the selection of the word it and for its semantic relationships.

Continuity is an aspect of word order, so only one of these dependencies is relevant to continuity: the one from it to raining simply doesn't count. In contrast, the dependency from very to great in *very with great difficulty is the only one that very has, so it has to be relevant to word order - but the phrase based on it is discontinuous. This is the essential difference between permitted and illegal discontinuity: the presence of alternative continuous phrases in the former but not in the latter. In It was raining the discontinuity of it .. raining doesn't matter, because it also has a dependency which doesn't involve discontinuity; but this is not true of very in *very with great difficulty.

How can we translate this idea into rules, principles and diagrams? Dependencies are shown in WG by labelled arrows that point towards the dependent, and every arrow must be specifically permitted by the grammar - the grammar must allow that kind of dependency between words that belong to those word classes and that are in that order. For present purposes we can make do with prose statements, so here is a little grammar-fragment that generates the dependency structure for with very great difficulty that is shown in **Figure 2**. (Much of this information will in fact be inherited from more general statements in the grammar, but for simplicity we can focus on the output of the inheritance.)

(30)   A grammar for with very great difficulty.
  a.   with is a preposition (P[12]).
  b.   very is an adverb (A).
  c.   great is an adjective (J).
  d.   difficulty is a noun (N).
  e.   with has a complement (c).
  f.   The complement of with is a noun.
  g.   The complement of with follows it.
  h.   The parent of very is an adjective or adverb.
  i.   very is an adjunct (label: `a') of its parent.
  j.   very precedes its parent.
  k.   The parent of great is a noun.
  l.   great is an adjunct of its parent.
  m.   great precedes its parent.

Unfortunately this grammar also generates a structure for *very with great difficulty as can be seen in **Figure 2**. It would be very difficult to rule out the latter example by changing the grammar, but it can be done very easily by means of a pair of general principles that require dependency structures to be both tangle-free and complete (free of `dangling' words that have no parent to dependend on), the No-tangling and No-dangling Principles.

(31)   **The No-tangling Principle**
        Dependency arrows must not tangle.
(32)   **The No-dangling Principle**
        At least one dependency arrow must point at every word.

with very great difficulty
P   A   J   N

*very with great difficulty
A   P   J   N

Figure 2

What about examples like <u>that it was raining</u>, where discontinuity (<u>it ..</u>
<u>raining</u>) is allowed? Once again it would be difficult to distinguish them by the
grammar rules themselves. Here is a little grammar for this example which
sanctions each of the dependencies in the first diagram of **Figure 3**. The term
`sharer' corresponds to the descriptive notion of a `coherent' complement, and to
the LFG term `XCOMP' and the earlier WG term `incomplement' (Hudson
1990:117). It is abbreviated `r' to avoid ambiguity with `subject'.

(33)   A grammar for <u>that it was raining</u>.
      a.   <u>it</u> is a noun (label: n).
      b.   <u>was</u> is a finite verb (v:f).
      c.   <u>raining</u> is a non-finite verb (V:n).
      d.   <u>was</u> has a subject (label on the arrow: s).
      e.   The subject of <u>was</u> is a noun.
      f.   The subject of <u>was</u> precedes it.
      g.   <u>was</u> has a `sharer'.
      h.   The sharer of <u>was</u> is a non-finite verb.
      i.   The sharer of <u>was</u> follows it.
      j.   (NB!) The subject of <u>was</u> is also the subject of its sharer.
      k.   <u>that</u> is unclassifiable (label: t[13]).
      l.   <u>that</u> has a complement.
      m.   The complement of <u>that</u> is a finite verb.
      n.   The complement of <u>that</u> follows it.

The grammar is crude, as any reader can appreciate, but it is sufficient to justify
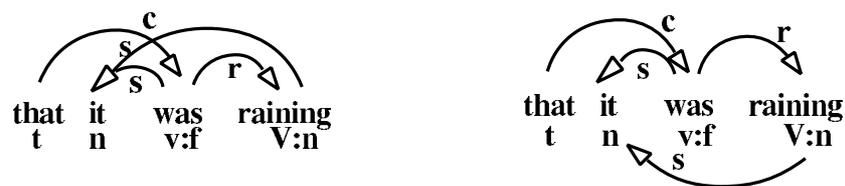every dependency in this diagram.

that   it   was   raining
t   n   v:f   V:n

that   it   was   raining
t   n   v:f   V:n

Figure 3

The second diagram in **Figure 3** is the crucial one because it shows why the
discontinuity is permitted. We can build a complete and tangle-free dependency
structure even without the one that causes tangling; this is the structure that is
drawn above the words, and for obvious reasons we can call it the `surface
structure'; the offending dependency is relegated to the basement. Every sentence
must have a surface structure in which every word depends on one other word,
and instead of banning all tangling we can ban it only in the surface structure. This
idea is expressed in the Surface Structure Principle (which is meant to replace the

Adjacency Principle of Hudson 1990:114; for discussion see Hudson 1994b, Rosta 1994).

(34)   The Surface Structure Principle
        Only a subset of a sentence's total dependencies (its `surface structure')
        must satisfy word-order constraints (including the No-tangling Principle) and
        the No-dangling Principle.

In short, surface dependency structure has to be translatable into a phrase structure, but anything goes among the remaining dependencies (provided, of course, that they are individually permitted by the grammar).

The Surface Structure Principle is one of a small set of mechanisms that WG offers for handling discontinuity, the others being specialised for handling coordination and clitics (Hudson 1990:118, 1994b). According to WG, therefore, any example of discontinuity which is not caused by coordination or cliticization must be analysable in terms of multiple dependencies in which a tangling dependency which shows a discontinuity is `supported' in the surface structure by another which does not. Muliple dependency in WG corresponds to structure-sharing in HPSG, which Pollard and Sag describe as the `central explanatory mechanism' of the theory, comparable with `move α' in GB theory (1994:19).

There is one small gap in the discussion which we must fill before moving on. The No-dangling Principle says that every word must have a supporting dependency arrow, but in each of our diagrams there is one word which has no arrow because it does not depend on any other word - the `root' of the phrase. The No-dangling Principle seems to be wrong. Moreover, so long as one word has no arrow, the No-tangling Principle won't give the desired results either. Take *great with difficulty, which we certainly want to exclude. The fact is that it is possible to draw a tangle-free structure as shown in the first diagram in **Figure 4**; but this is the wrong outcome. To prevent it we shall adopt a convention of allowing just one word to have a `virtual' parent, shown as a vertical arrow which is assumed to be infinitely long. This automatically tangles with any offending arrow as shown in the second diagram. The third diagram shows the effect of adding this arrow to a full sentence, It is raining.
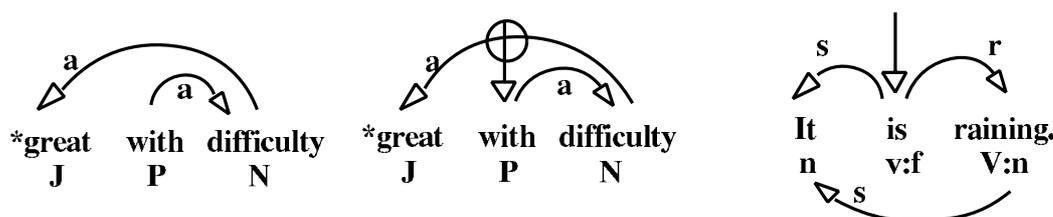

Figure 4

There is more to say about the general principles that control surface structure but we can leave it till section 7. The parts of WG theory that I have explained so far are already enough to take us through the next two sections, where we shall consider how WG can handle all of German PVPF except for the most difficult part, co-fronted subjects.

5.   V2 and simple fronting
The first fact that we recorded in our summary (25) is that "a declarative non-

dependent finite verb must have just one preceding dependent, its `X'''; this is the well-known V2 requirement[14] expressed in terms of dependencies. The complication, of course, is that the verb's X is not just its X; it is also its subject, object, adjunct or whatever - or indeed a dependent of a subordinate verb, thanks to extraction. For example, in <u>Er schläft</u>, `He sleeps', `he' has two different dependency relations to `sleeps': X and subject. We already know that a word may have two different parents, so it is hardly surprising that it may have two different relationships to a single parent. In principle we could draw a separate arrow for each relationship, but it is simpler to add two labels to one arrow: `s,x'. (I have reduced `X' to lower-case `x' to bring it into line with the other arrow labels.) **Figure 5** is the diagram for this simple example. This analysis is the dependency analog of a standard GB analysis (Haegeman 1991:528) in which the finite verb is raised to C and X is raised to Spec of C.

s,x

Er   schläft.
n   V:f
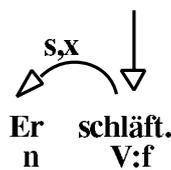
Figure 5

The grammar rules are easy enough to state in prose:
(35)   Simple fronting
      a.   A finite verb which has a virtual parent has one X.
      b.   A verb's X also has some other dependency relationship to it.
These rules allow one dependency to be labelled `x' and require `x' to be combined with some other label. In themselves they say nothing about word order, so they have to be supplemented by word-order rules. The following rules may not be right, but they are a good first approximation.
(36)   Word order
      a.   A word follows its dependents.
      b.   A finite verb which has a virtual parent precedes its dependents.
      c.   A finite verb which has a virtual parent follows its X.
The application of these rules is controlled by default inheritance, which automatically[15] gives priority to the specific over the general. Rule (a) assumes that German, unlike English, is basically a head-final language. Rule (b) overrides this in the case of the main finite verb, and rule (c) overrides this override just in the case of this verb's X dependent.

One of the arguments often used in favour of phrase structure is that it reflects the `structure-dependent' nature of word-order rules: what the rules control is usually not a single word, but a whole phrase. For example, subject-verb inversion does not invert just the subject noun, but its whole phrase. How can a grammar without phrase structure be structure-dependent in this sense? In case the answer is not already obvious, it involves the No-tangling Principle. Suppose word W1 depends on W2, and some word-order rule requires W1 to precede W2. The No-tangling Principle guarantees that all the dependents of W1 (and their dependents, recursively) also precede W2, because otherwise the phrase rooted in W1 will be split by W2 and dependencies will tangle in the surface structure. In

16

general the effect of this principle is to guarantee the same integrity of phrases as PS.

Take the `first position' in a German main clause, for example. In PS terms this is filled by a complete phrase, but in dependency terms it is filled just by a single word, the finite verb's X; but No-tangling guarantees that this word brings its whole phrase with it, so to speak. If any of the X's dependents `stayed behind' to the right of the finite verb, surface tangling and ungrammaticality would ensue. Consider the examples in 37):

(37) a. [Man]$_X$ findet viele Museen in dieser Stadt.
         one finds many museums in this town
      b. [In dieser Stadt]$_X$ findet man viele Museen.
      c. *[In]$_X$ findet man viele Museen [dieser Stadt].
      d. *[In dieser]$_X$ findet man viele Museen [Stadt].

The first two sentences are unproblematic because X's phrase is continuous: either it has no dependents (a) or its dependents are clustered round it (b). In the other two sentences, in contrast, the dependencies within the X phrase tangle with those in the rest of the sentence, as shown in **Figure 6**. (For simplicity I shall omit irrelevant labels from now on.)
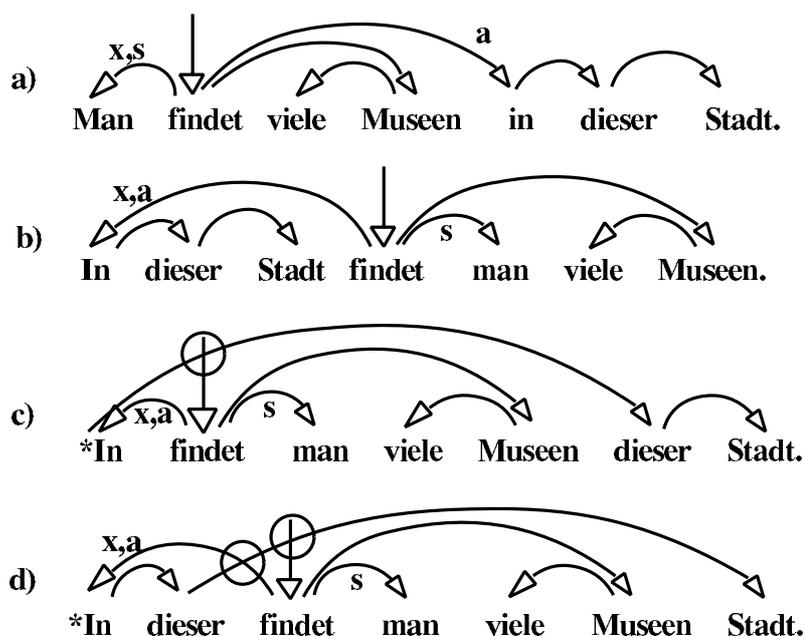

Figure 6

A strict application of the No-tangling Principle gives just the right results. Notice that in this example the Surface-structure Principle is irrelevant because there is nothing but surface structure; each word has just one parent, and in particular dieser and Stadt do not depend on the finite verb, so they can't take their position from it.

The analysis so far leaves us with precisely the same problem as a phrase-structure analysis has: how to explain PVPF. How is it that if X is a non-finite verb it does **not** have to carry its whole phrase with it? This is our topic in the next section.

6. Partial-VP fronting and generalized raising

The second fact in our summary of German PVPF is that `if X is a non-finite verb, any of its dependents may either be co-fronted with it or follow the finite verb.' Let's take the same simple examples that we quoted at the start of the paper:

(38) a. $[Er]_X$ darf Kaffee trinken. `He may drink coffee.'
      b. $[Kaffee\ trinken]_X$ darf er.
      c. $[Trinken]_X$ darf er Kaffee.

The first two examples are easy to explain on standard assumptions: `may' is a subject-to-subject raising verb, so `he' is subject of `drink' as well as of `may'; `drink' is complement (`sharer', abbreviated to `r') of `may'; and the word-order rules are those given in (36). On the same assumptions, example (c) should be ungrammatical because `coffee' depends only on `drink', but causes tangling by taking its position from `may'. The relevant structures are shown in **Figure 7**.
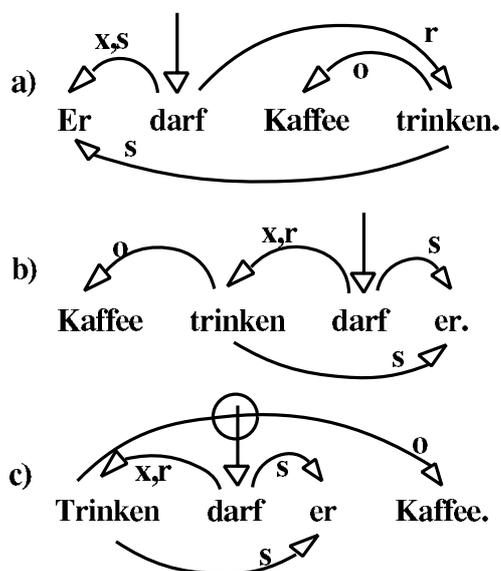
Figure 7

      As I explained in section 3, the main idea running through the HPSG analyses of PVPF is what is variously called `clause union' or `argument attraction', which involves structure-sharing of the un-fronted dependents of X. Example (38c) is grammatical because `coffee' is structure-shared by both verbs, rather than belonging exclusively to `drink'. I suggested the name `generalized raising' because of the similarity to subject-to-subject raising.

      How, then, do we arrange for generalized raising in just the right cases? It must be the result of a parochial rule of German, because the corresponding English sentences are ungrammatical; and it must be a peculiarity of non-finite verbs, because fronted prepositions (for example) do not allow it, as we saw in (37). It must be optional, because each dependent of X is free to raise or not, regardless of the other dependents. Here is a rule which satisfies all these requirements, and which illustrates the benefits of being able to operate on individual dependencies rather than whole phrases; we shall revise it slightly below.

(39)   Generalized raising (first version)
      If W1 is sharer of W2, any dependent of W1 may also be a dependent of

W2.

This is the foundation for our explanation of PVPF, and plays a role similar to that of scrambling and structure-sharing in the GB/MP and HPSG analyses.

It applies to any pair of verbs W1 - W2 in the `sharer' relationship, which in more standard terminology is a `coherent' relationship. It applies freely to any dependent of W1, whether subject, complement or adjunct, and allows it to have W2 as an additional parent; but this is a decision which is made separately for each dependent, so some dependents may be shared with W2 while others remain simply as dependents of W1. If W1 is fronted, its dependents will be treated differently according to whether or not they are shared with W2: those that are, take their position from W2, but the remainder co-front with W1. At this stage it does not matter what we call the added dependency; I shall simply leave it unlabelled in diagrams. With this rule in place we can give a satisfactory analysis for the grammatical <u>Trinken darf er Kaffee</u>; this is in **Figure 8**.
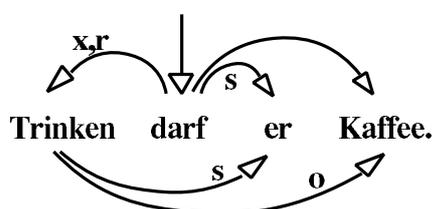

Figure 8

The effects of generalized raising are easy to see in PVPF. What about non-fronted examples, where Generalized raising is also available? In some cases the effects of raising are visible here too. The evidence for raising comes from examples like the following (from Grewendorf and Sternefeld 1990:9 and Haider et al 1995:9).

(40)   a.   weil den Max$_{Acc}$ jeder$_{Nom}$ zu kennen glaubt
            because the Max everyone to know believes
            `because everyone thinks they know Max'
       b.   *weil den Max jeder zu kennen bedauert
            because the Max everyone to know regrets
       c.   daß sie ihn nicht zu stören wagt
            that she him not to disturb dares
            `that she dare not disturb him'

The only way to explain the order of elements in example (a) is to assume that the object of `know' also depends on `believes'; on that assumption `Max' and `everyone' both share the same surface parent so they can occur in either order. This is possible because `know' is sharer of `believes', in contrast with example (b) where `know' is the `incoherent' complement. Similarly for example (c), where `him' must be raised to depend on `dares' because `not' separates it from its other parent, `disturb'.

On the other hand, there is also evidence that generalized raising is optional in non-fronted examples, just as it is with fronting. The next example (from Haider et al 1995:11) shows this clearly, because `the (story)' must depend only on `read', in contrast with `to a (child)' which must depend on `may' as well as on `read'.

(41)   wenn er einem Kind hätte das Märchen vorlesen dürfen

if he to.a child had the fairy.tale read.to may
`if he had been able to read the fairy tale to a child'

The structure for this example is shown in the first diagram of **Figure 9**. The price we pay for allowing examples like this is spurious ambiguity in more straightforward examples like <u>wenn er das Märchen vorlesen darf</u>, `if he may read the fairy tale'. The other diagrams show the alternatives for this sentence, but we should note that spurious ambiguity is widespread and need not be seen as an analytical weakness.
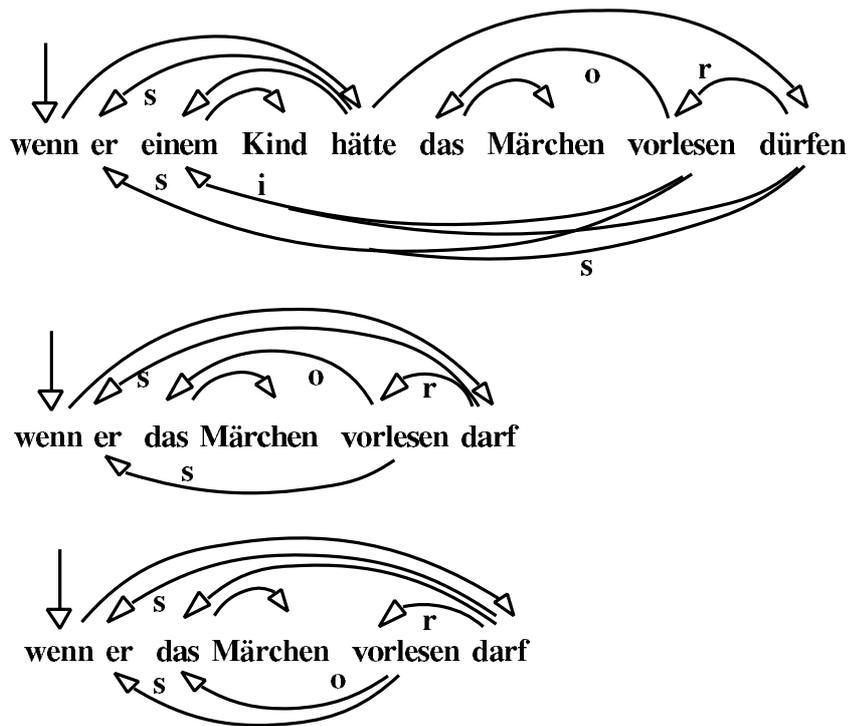
Figure 9

Let's move on to the more complicated cases covered by our third generalization: `if another non-finite verb depends on a non-finite X, its dependents may either be co-fronted or follow the finite verb just like those of X itself'. In other words, a limited amount of recursion is built into generalized raising. Here is a relevant example (= 6c):

(42)   [Geben können]$_X$ wird Peter Maria das Buch.
         give can will Peter Mary the book
         `Peter will be able to give Mary the book.'

Suppose this example represents the upper limit of recursion; we can then accommodate it by a slight revision to the Generalized Raising rule:

(43)   Generalized raising (final version)
         If W1 is X of W2, any dependent of (a sharer of) W1 may also be a
         dependent of W2.

Since `give' is the sharer of `can' its dependents (`Mary' and `the book') can be raised to depend on `will' as though they were dependents of `can' itself. If recursion is freer than I am assuming, the rule will need to be changed following the model of subject-to-subject raising, which is completely recursive: `Peter' is the

subject of `can' because it is the subject of `will', and of `give' because it is subject of `can'. Maybe a similarly recursive rule is needed for the other dependents, but I must leave the question open for lack of data.

This double-infinitive construction leads directly to our fourth generalization. `If one non-finite verb (V1) is the complement of another (V2), V2 must immediately follow V1.' (We can now understand `complement' to mean `sharer'.) This generalization was meant to account for the impossibility of examples like the following:

(44) a. *[Müssen]$_X$ wird er seiner Tochter ein Märchen erzählen.
           must will he to.his daughter a fairy-tale tell
           `He will have to tell his daughter a fairy-tale.'

      b. [Erzählen müssen]$_X$ wird er seiner Tochter ein Märchen.

`Tell' and `must' are both non-finite, and `tell' is the sharer of `must' so the only possible order is `tell must', without any intervening words. Given the general freedom of generalized raising, this restriction is surprising. How can we express it in WG, and (more important still) how can we explain it?

One possible explanation runs as follows. What is special about these double-verb combinations is that they involve subject-raising: V1's subject is raised to double up as V2's subject as well. This is `raising' because V2 is higher than V1 in terms of the usual metaphor applied to dependencies in which a word is `lower' than the word it depends on. But this raising must be genuine - it must show up in the surface structure, where V1 must depend on V2. This is where `height' can be measured without contradiction, in contrast with the non-surface dependencies where all sorts of contradictory possibilities are allowed (including mutual dependency). But the bad examples are bad precisely because V1 does not depend on V2 in the surface structure. Consider these simple examples:

(45) a. [Er]$_X$ wird schwimmen können.
           he will swim can
           `He will be able to swim.'

      b. [Schwimmen können]$_X$ wird er.

      c. *[Können]$_X$ wird er schwimmen.

The structures for these three sentences are shown in **Figure 10**, with the irrelevant arrows dotted. In the two good sentences, `swim' has `can' as its surface parent, so `he' is genuinely raised from a lower word to a higher one. In contrast, `swim' still depends on `can' in (c), but the dependency is excluded from the surface structure so `swim' cannot share its subject with `can'.
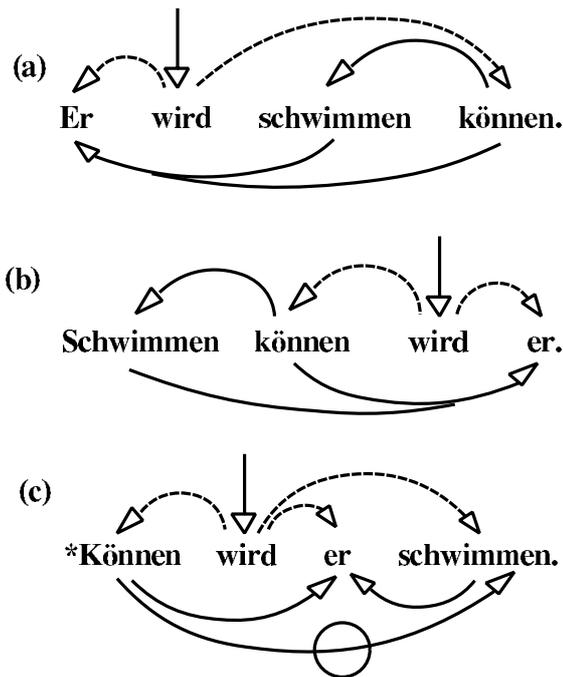
Figure 10

To formalise this explanation we need a new principle, the Raising Principle, which will play a crucial part in the discussion of co-fronted subjects. The Raising Principle controls the dependency relationships among the three elements involved in raising: the shared word D (for dependent) and its two parents Pp and Pd, where Pp is the parent of Pd. (In the last example, `he' is D, `swim' is Pd and `can' is Pp.) This principle is universal (though I shall say in section 7 that it can be overridden in specific constructions and languages).

(46)   The Raising Principle (first version)
         If D is a dependent of both Pp and Pd,
              and Pd is a dependent of Pp,
         then Pp must be the surface parent of Pd.

This principle rules out the bad structure in **Figure 10** because `can' is not the surface paent of `swim', so these two words cannot share `he' as a dependent.

The same principle explains the badness of sentences in which a pair of clause-final infinitives are separated, as in (12) repeated below.

(47)   a.   *Ich wusste, daß er das Examen bestehen würde können.
                  I knew that he the exam pass would can
                  `I knew that he would be able to pass the exam.'
         b.   Ich wusste, daß er das Examen bestehen können würde.
         c.   Ich wusste, daß er das Examen würde bestehen können.

In all these sentences, `pass' depends on `can' and `he' is their shared subject. This relationship is controlled by the Raising Principle, which is satisfied in the good sentences (where `pass' is immediately in front of `can') but not in the bad one, where the surface parent of `pass' is `would', not `can'.

What the Raising Principle restricts so far is the relationship between Pp and Pd, which must be a surface dependency - i.e. one that is `visible' (and presumably usable by a parser). But raising also means that the raised word D forsakes its

22

lower parent Pd and clings visibly to Pp - i.e. that D takes its position from Pp rather than from Pd. Nothing that we have said so far guarantees this, and as things stand nothing forces a grammar to treat Pp rather than Pd as D's surface parent. This matters in pairs like the following:

(48) a.   Pat kept talking.
      b.   *Kept Pat talking.

Why can't we generate both of these using the same grammar, but with different surface parents for <u>Pat</u>: <u>kept</u> in (a) and <u>talking</u> in (b), as shown in diagram (b1) of **Figure 11**?
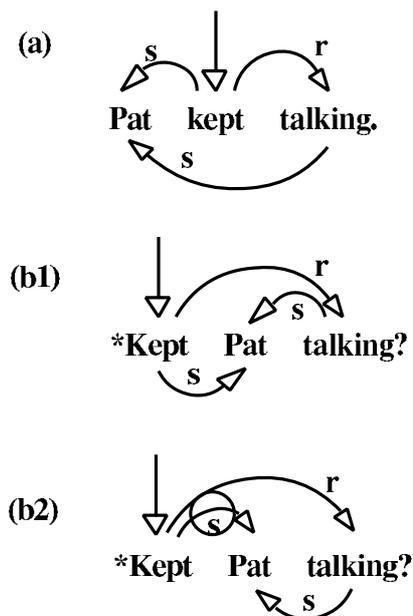

(a) Pat kept talking.

(b1) *Kept Pat talking?

(b2) *Kept Pat talking?

Figure 11

It's true that English requires non-auxiliary verbs to follow their subjects, but this is a word-order rule and (by the Surface-structure Principle in (34)) word-order rules only apply to surface structure. If <u>kept</u> has to be the surface parent of <u>Pat</u> as in (b2), this rule must apply and ungrammaticality is predicted; but if it could be <u>talking</u> (b1), the rule can be avoided and the grammar wrongly generates (b). This is clearly wrong, so we must strengthen the Raising Principle:

(49) The Raising Principle (final version)
         If D is a dependent of both Pp and Pd
              and Pd is a dependent of Pp,
         then Pp must be the surface parent of Pd
              and Pd must not be the surface parent of D.

(The Chomskian analog of this version of the principle is the Empty-Category Principle, which requires traces to be c-commanded by their antecedents; but this also has a range of other effects that go beyond the Raising Principle. A similar principle is implicit in HPSG, where it is taken for granted that a structure-shared element belongs in the higher structure.)

         The revised Raising Principle explains why a shared subject has to be raised so that it takes its position from its higher parent. It also explains something that we took for granted in the earlier discussion of generalized raising: why the additional

dependency link to the higher verb forced the dependent to follow it. Take our earlier example.

(50)   [Trinken]$_X$ darf er Kaffee.
        drink may he coffee

We explained the non-fronting of `coffee' by making it depend on `may' as well as on `drink'; but this explanation only works if a dependent has to take its position from its higher parent. This is guaranteed by the revised Raising Principle.

We now have the makings of a grammar which will generate grammatical partially-fronted VPs and which excludes those which are ungrammatical. We are left with the hardest part, the co-fronting of subjects.

7.   Co-fronted subjects

Our last fact about German fronting was a complex one, but we can divide it into two parts: the possibility of co-fronting a subject, and the restrictions on doing so. The main problem is not to explain these restrictions but to explain why subjects can be co-fronted at all, and this is our topic for the present section. Take one of our simpler examples:

(51)   [Eine Concorde gelandet]$_X$ ist hier noch nie.
        A Concorde landed is here still never
        `A Concorde has never yet landed here.'

The trouble is that this sentence has just the kind of structure that we have just ruled out through the revised Raising Principle. The verbs `landed' and `is' share their subject, `a (Concorde)', but this takes its position from the lower of them, `landed' as shown in **Figure 12**. In contrast, its paraphrase in (52) avoids the problem.

(52)   [Eine Concorde]$_X$ ist hier noch nie gelandet.

In short, co-fronted subjects are lowered so they fall foul of the Raising Principle (and, in Chomskian terms, they conflict with the Empty-Category Principle); so how can they be grammatical?
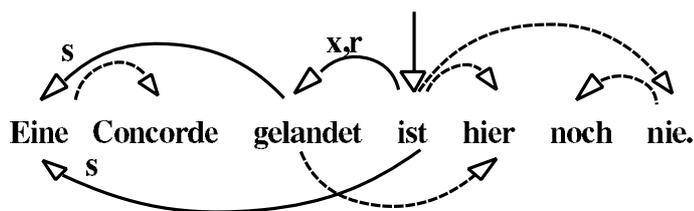

Figure 12

What we must bear in mind is that co-fronted subjects are a peculiarity of German and Dutch, so we want Universal Grammar to continue to ban them in general. It is possible, of course, that they reflect some more general characteristic of these two languages but no candidates are currently on offer. (Even if it is possible to relate PVPF to general freedom of word order in the clause, this will not help in explaining the lowering of subjects.) It would be intellectually satisfying to explain co-fronted subjects in terms of some deep parametric difference, but by no means essential because there are easy alternatives. Why are co-fronted subjects allowed at all in any languages? Because they increase the communicative options by allowing users to define the topic in terms of the action and its subject-

participant. Why in German and Dutch? Because these two languages already allow a range of topicalisation options which is otherwise unrestricted, thanks to PVPF, so thus fills the remaining gap. How do children learn that subjects can be co-fronted? By hearing them.

When such simple explanations are available, why should we look beyond the obvious `brute-force' solution. German has a rule for fronted sharers which overrides the general Raising Principle (by default inheritance). Overriding requires strict parallellism in the facts concerned so it is important that the overriding fact should have the same logical form as the overridden one in (49).

(53)   Subject co-fronting
        If D is the subject of Pp and Pd,
            and Pd is both X and sharer of Pp,
        then Pp must be the surface parent of Pd
            but Pd may be the surface parent of D.

In WG there is no theoretical difference between rules and principles, so by default inheritance this rule can take priority over the general principle. It allows a shared subject to take its position from the lower of its parents, thereby allowing co-fronting. In section 8 we shall see that similar rules for overriding the Raising Principle exist in French and Icelandic. The following section will return to the restrictions on co-fronted subjects.

8.   Lowering in other languages

The Raising Principle has the effect of banning `lowering', a pattern of dependencies in which a word takes its position from the lower of two parents. We have just recognised an exception to the principle in German, so we should look for precedents in other languages. The data are familiar, as is the idea of some kind of `subject-lowering' operation, but it is worth reviewing the evidence from a dependency perspective. This will also give me a chance to show how WG can accommodate differences between languages, because we shall see that although each of these exceptions to the Raising Principle involves subjects, the details of the exception are different in each case.

We start with `stylistic inversion' in French (Kayne and Pollock 1978, Roberts 1993:216-9). which is found (among other constructions) in relative clauses like the following[16]:

(54)   a.   le livre que Pierre voulait encore emprunter à son père
                the book which Peter wanted again borrow from his father
                `the book which Peter wanted to borrow again from his father'
        b.   le livre que voulait encore emprunter Pierre à son père

The order in (a) is the one predicted by the Raising Principle, because `Peter' is shared by both `wanted' and `borrow', but takes its position from the higher of them. Example (b) is the problematic one, because `Peter' takes its position from `borrow'. The important fact about this example is that `Peter' separates `borrow' from one of its dependents `from (his father)', so it too must have a surface dependency on `borrow'. (In GB/MP terms, it remains in its basic position within the final VP.) But at the same time there must be a dependency between `Peter' and the finite verb `wanted' because of their inflectional agreement; in at least the WG version of dependency analysis, agreement is always carried by a dependency between the words concerned. The WG structure for (b) given in **Figure 13** (with

irrelevant arrows dotted) shows the `lowering' effect of this double dependency.
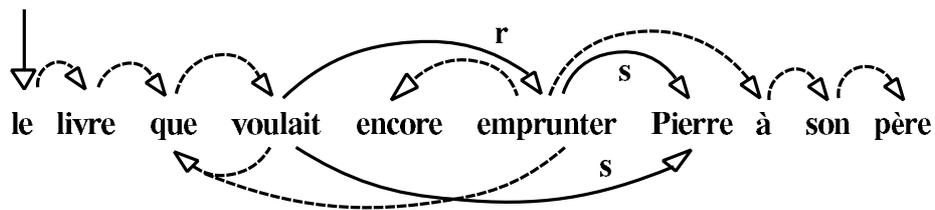

Figure 13

What rule would override the Raising Principle in these French examples? What follows is only a first approximation, which will certainly need to be revised in order to allow stylistic inversion in other constructions, as well as to allow recursion, but the changes will only affect the part of the rule that restricts it to the first two verbs in a relative clause. In dependency terms, a French relative clause may be defined by a finite verb (`wanted' in our example) which depends on a relative pronoun. This verb is Pp in the rule, and its lowered subject is D.

(55)   French relative-clause subjects
           If D is the subject of Pp and Pd,
                and Pd is sharer of Pp,
                and Pp depends on a relative pronoun,
           then Pp must be the surface parent of Pd
                but Pd may be the surface parent of D.

Notice the great similarity between this rule and our rule for German fronted subjects (53). No doubt rather similar rules are needed for other Romance languages that allow free inversion.

Icelandic seems to allow a somewhat different kind of subject-lowering. It is generally agreed that the derived subject (and underlying object) of a passive participle may optionally follow it, leaving the top surface subject position to be filled by an expletive það, `it/there'. The examples in (56) are from Andrews (1990:191-2), and those in (57) from Sigurðsson (1991:346).

(56)   a.   [Einhver leikari$_{Nom}$] var kosinn í forsetaembættið.
                 some actor was elected to the.presidency
           b.   það var kosinn [einhver leikari$_{Nom}$] í forsetaembættið.

(57)   a.   það höfðu$_{plur}$ sennilega sokkið [einhver bátar$_{Nom}$] í firðinum.
                 there had probably sunk some boats in the bay
           b.   það mundu$_{plur}$ kannski hafa verið seldir$_{Nom\ plur}$ [einhverjir bátar$_{Nom\ plur}$] á uppboðinu.
                 there would perhaps have been sold some boats at the.auction

It is quite uncontroversial to assume that the bracketed phrase in each example depends on both the first and the last verb in each example for the following reasons. Starting with the first verb, the bracketed phrase must be structurally related to the finite verb (var, höfðu or mundu) because it determines the latter's singular/plural agreement according to the general rules for subject-verb agreement. Any syntactic rule that links two words implies a dependency between them, so one must depend on the other and the easiest assumption is that the noun depends on the verb as some kind of subject. In Icelandic the typical `subject' properties can be divided between two dependents and there are well-known

reasons for recognising some of them in það; but WG allows us to subclassify dependency relationships, so we can split `subject' into two sub-types called `first-subject' and `second-subject' assigned respectively to það and to the full noun. In (56a), `some (actor)' is the full subject of `was' whereas in (56b) it is just its second-subject. But the important point to establish is the existence of some dependency to the first and higher verb.

On the other hand, `some (actor)' in (56b) must also depend on the lower verb `elected', and it is from this verb, not the higher one, that it takes its position. This can be seen clearly from the fact that it is followed by `to (the presidency)', which must depend on `elected'. As Andrews points out, if `some (actor)' separates `elected' and a dependent of `elected', it must itself be a surface dependent of `elected' (rather than simply a delayed subject of `was'). (In GB/MP terms, it remains in situ within the lower VP.) Similarly for the examples in (57), where Sigurðsson takes it for granted that the nominative NP is inside the last VP. A `lowering' analysis in which the nominative NP takes its position from the lower of its two parents is hard to avoid, though it is contrary to the Raising Principle.

This pattern of subject-lowering is not restricted to passives, but is also found as an alternative to raising. The following is from Andrews (1990:192).

(58) a.   Einhver leikari ætlar að verða kosinn í forsetaembættið.
           some actor `intends' to be elected to the.presidency
           `Some actor seems likely to be elected president.'
     b.   það ætlar að verða kosinn einhver leikari í forsetaembættið.

Once again the position of `some (actor)' between the lower verb and its other dependent forces us to take it as a surface dependent of this verb. The rule for Icelandic may be as follows:

(59)   Icelandic lowered subjects
        If D is the second-subject of Pp and Pd,
            and Pd is sharer of Pp,
            and D is the object of Pd,
        then Pp must be the surface parent of Pd
            but Pd may be the surface parent of D.

This rule can probably apply recursively to explain virtuoso examples like the following (Andrews 1990:211-2):

(60) a.   Honum$_D$ er talið hafa verið gefnir$_N$ peningarnir$_N$.
           to.him is thought to.have been given money
           `He is thought to have been given money.'
     b.   Honum$_D$ eru taldir$_N$ hafa verið gefnir$_N$ peningarnir$_N$.
           to.him are thought to.have been given money

It is not certain that `money' is a surface dependent of `give', but if it is, it is lowered into this position in sentence (b) but not in (a). What makes both examples complicated is that this lowering of the second-subject is combined with raising of the first-subject. Tentative WG analyses are offered in **Figure 14**.
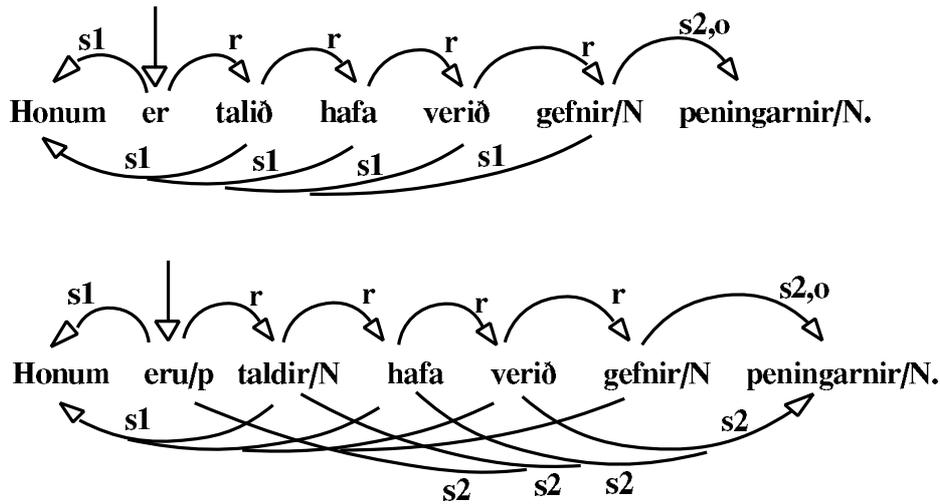
Figure 14

In conclusion, there is clear evidence of subject-lowering in French (and other Romance languages) and Icelandic. At an abstract level there are just minimal differences among the rules for these languages and for German, but the surface difference is striking: subject delay in French and Icelandic contrasting with subject co-fronting in German. The main point at issue is simply whether lowering is possible in principle, and we have established quite clearly that it is. This conclusion establishes an important theoretical finding: that even universal principles apply by default inheritance and may be overridden. But for our immediate purposes the main conclusion is simply that we can feel reasonably confident that subject-lowering is the right analysis for co-fronted subjects. Now that we understand why subjects can be co-fronted at all, we can move on to the restrictions which sometimes prevent a subject from being co-fronted.

9.   Restrictions on co-fronted subjects

We turn now to the restrictions on co-fronted subjects that we identified in section 2.

(61)   The subject of a non-finite X (which is also the subject of the finite verb) may be co-fronted provided that:
      a.   it is indefinite, and
      b.   it has low transitivity - unaccusative, passive or free of non-pronoun complements.

The definiteness requirement could be added to the Subject co-fronting rule (`If D is indefinite and ...'), but it may have a pragmatic explanation. I leave this open. The low-transitivity restriction is more challenging.

The following explanation is more tentative than the rest of the paper, but it is worth offering as an avenue that is worth exploring. It involves the interface between syntax and semantics, at the point where X has to be mapped onto a coherent meaning. The idea is that some co-fronted subjects can combine with the fronted verb to produce a meaning which survives the additions required by the rest of the sentence, while others cannot. The idea of a Concorde landing (52) is compatible with the rest of the sentence, but that of an outsider winning (24a) is

28

not. I shall first flesh out the general idea a little, and then show how it translates into more formal WG structures.

The first thing to establish is that X must have a meaning - i.e. it must correspond to some identifiable part of the semantic structure. The choice of X is not determined by its grammatical or even semantic role - any dependent of the verb, expressing any semantic role, will do. How do speakers choose which part of a sentence will be its X? X is standardly defined as the clause's topic (Durrell 1996:456), the element whose meaning is taken as the `starting-point' for the rest of the clause's meaning. Therefore X must have a meaning, and every distinct X must have a distinct meaning. For example, `eaten' and `eaten an apple' must have distinct meanings:

(62)　a.　　[Gegessen]$_X$ hat er einen Apfel.

　　　　　　eaten has he an apple

　　　b.　　[Einen Apfel gegessen]$_X$ hat er.

Sentence (a) is about eating, while (b) is about eating an apple, and similarly for all the multiple alternatives with fronted partial VP illustrated in our earlier examples. (See especially the examples in (4).)

This does not mean that the overall meaning of the sentence varies with the choice of X; on the contrary, it seems reasonable to assume that the sentences, as such, generally have the same meaning regardless of the choice of X (with exceptions like the one mentioned in the next paragraph). But this meaning is reached compositionally via different routes. Assuming - as we surely must - an incremental parser which builds the semantic and syntactic structures in parallel as the words are processed, the combination of `eat' and `an apple' is achieved at different moments in the processing of examples (a) and (b) - by the end of X in (b), but not till the end of the sentence in (a). The main conclusion, then, is that X must have a meaning, and that this meaning must be compatible with the various additions which will be provided by later words.

In case this conclusion needs further support, we can mention an interesting observation by Netter (1991:33) about the interpretation of modal verbs. He quotes the following examples:

(63)　a.　　[Ein Häftling entspringen]$_X$ muß hier öfters.

　　　　　　a prisoner escape must here frequently

　　　　　　`It must frequently be the case here that a prisoner escapes.'

　　　b.　　[Ein Häftling entspringen]$_X$ darf hier nicht.

　　　　　　a prisoner escape may here not

　　　　　　`It must not be the case here that a prisoner escapes.'

In both sentences the modality applies to the content of X, `that a prisoner escapes', and not to the notion of a prisoner. He excludes the interpretations `A prisoner frequently has to escape here' and `A prisoner does not have permission to escape here.' This difference is as we might expect if X defines a meaning in the total semantic structure; but it is also an interesting example where the choice of X affects the final meaning given to the whole sentence.

Rather obviously, it also follows that what follows X must be worth saying. There would be no point in packing the whole sentence (except the finite auxiliary) into X, because the sentence would say nothing about its defined topic. Hence the impossibility of sentences like (a) below.

(64)　a.　　*[Er einen Apfel gegessen]$_X$ hat.

b. *[Er gegessen]$_X$ hat einen Apfel.

But what's wrong with example (b)? Why can't we use this as an answer to a question about what he has eaten? At this point pragmatic explanations run out and we look to the grammar.

Suppose there is a grammatical rule or principle that requires the subject to close off the verb's semantics; once the subject's meaning has been added, nothing else may be changed. (This implies a hierarchy on which the subject stands at one extreme, such as the familiar Accessibility Hierarchy of Keenan and Comrie 1977 or the Obliqueness Hierarchy of Pollard and Sag 1994.) Let's call this the **Closing-Subject Principle**. We shall make it more precise in the next section, but it is probably enough to think of the subject as `freezing' the verb's meaning so that it cannot be further modified[17]. This would explain the badness of (64b): the processor must recognise a meaning for X, but at this point in the processing there is no object to bind the second argument of `eat', the eat-ee; therefore it has to take the interpretation in which the eat-ee is indefinite `something'. The meaning of X is therefore `he eat something'. The Closing-Subject Principle freezes this meaning, so that when `an apple' is processed it cannot be integrated semantically. Similar explanations are possible for the other examples surveyed in section 2 in which a fronted transitive verb is ungrammatical because it has a following independent object (see examples (14b) and (24)).

This principle, then, allows subject co-fronting with intransitive verbs, whether ergative (unaccusative) or not. (According to Haider (1990:94) Dutch allows it only with ergative verbs, which implies a constraint which we shall not need for German[18].) But non-ergative verbs only allow subject co-fronting provided they are what I called `complete', which means that if they have an overt object outside X, it must be a pronoun attached to the finite verb as in the following:

(65)   [Ein Außenseiter gewonnen]$_X$ hat das da noch nie.
         an outsider won has that there still never
         `An outsider has never yet won (that race).'

This sentence is grammatical with or without the object pronoun `that', though it is ungrammatical if the object is a full NP such as `the Derby'. I can only speculate about the explanation, but there must be some reason why a pronoun does not `count' as a non-fronted argument. Maybe it is simply a processing matter: the pronoun reaches the processor so soon after the end of X that it can be integrated into X before its interpretation is `frozen' by the subject. Alternatively, maybe there is a more formal reason to do with the positioning of pronouns, since (as noted earlier) they tend strongly to cluster after the finite verb. If this is a natural position even for pronouns that `really' belong in X, the problem would be solved. Unfortunately I cannot push these speculations further.

In contrast with unergative verbs, ergative and passive verbs are free to have unfronted arguments regardless of whether or not these are pronouns. Here are two of the relevant examples (Haider 1990:97):

(66)   a.   [Ein Fehler unterlaufen]$_X$ ist auch schon mal diesem Professor.
               a mistake happened is also already once to.this professor
               `This professor did in fact once make a mistake.'
        b.   [Ein solch schönes Geschenk gemacht]$_X$ wurde mir noch nie.
               a such beautiful present made was to.me still never
               `Such a beautiful present was never yet given to me.'

30

This suggests that the co-fronted subject does not have its usual closing-off effect, and it is tempting to link this freedom to the assumption that in both cases the subject is also the object of the fronted verb. Maybe its relevant feature is not its subject-hood but its object-hood? It is widely accepted that a verb's semantic bond to its object is especially close (Tomlin 1986 calls it `Verb-Object Bonding'), which implies another hierarchy of dependents, a `Verb-bonding hierarchy'. On this hierarchy the object is at one extreme, unlike the Accessibility and Obliqueness hierarchies which I mentioned above where the object is in the middle.

I propose that this hierarchy defines a second principle that overrides the first: the **Object-first Principle**. The verb and its object together define a meaning on which all other dependents build, so the object leaves the verb's meaning open to further changes. This would explain why the last pair of examples are both grammatical, because in both cases X's meaning is defined by the object and verb so some other arguments can be left open and available for binding by later dependents. These suggestions may provide a basis for understanding why verbs with `low transitivity' allow co-fronted subjects, as I shall show in the next section.

10.   The low-transitivity restriction on co-fronted subjects: a WG analysis
We start with a brief summary of the non-standard semantics used in WG (see Hudson 1984:131-210, 1990:123-65 and 1995a for further details). WG analysis produces two parallel structures: a syntactic structure based on word-word dependencies and a semantic structure in which meanings (senses and referents) are provided for the words in the syntactic structure. These various meanings are linked by dependencies and identities that may be predicted from the syntactic structure plus the entries for the words concerned. Consider the simple example <u>a picture of a girl</u>, whose structure is shown in **Figure 15**.
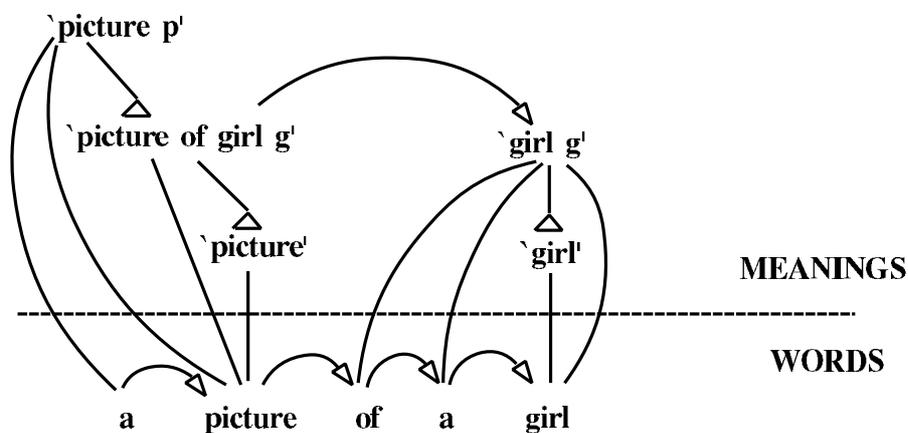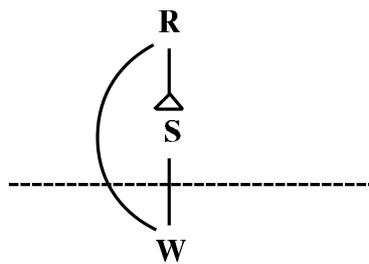

Figure 15

In this diagram the straight vertical lines relate words to their senses and the vertical arcs relate them to their referents[19]. The line leading to a little triangle shows the `isa' relation between an instance and its supercategory[20]. The notation is summarised in **Figure 16**.

```
          R
          |
          △
          S

------------|------------------

          W
```

KEY

S is the sense of word W.
R is the referent of word W.
R isa S.

Figure 16

What the analysis in **Figure 15** shows is that the sense of this example of the word <u>picture</u> - i.e. <u>picture</u> combined with <u>of a girl</u> - is derived compositionally from the sense of <u>picture</u> as found in the grammar - `picture' - plus the referent of its dependent. This is labelled `girl g', so the sense of <u>picture</u> is labelled `picture of girl g'. Assuming that the word is used to refer to some particular picture, the relationship between its referent and its sense is the familiar `isa' relationship between an instance and its superconcept, so `picture p', the picture concerned, isa the more general concept `picture of girl g'; and similarly the referent of <u>girl</u>, `girl g', isa the more general concept `girl'. As for the other two words, <u>of</u> and <u>a</u>, they both share the referent of their syntactic complements; so the first <u>a</u> is coreferential with <u>picture</u> while <u>of</u> and the second <u>a</u> are both coreferential with <u>girl</u>. One of the advantages of the WG approach to semantic analysis is the close integration of the syntactic and semantic dependencies: every syntactic dependency corresponds either to a semantic dependency or to a semantic identity.

The main point to notice in this example is the distinction between the three different kinds of meaning for <u>picture</u> and <u>girl</u>: the word's purely lexical sense (the sense stored in the grammar), the sense that is built compositionally, taking account of the word's dependents, and the referent, which isa the compositional sense. Each meaning is represented separately in the network, and each has a distinct name, though they are also linked in an isa hierarchy: the referent isa the compositional sense, which isa the lexical sense. Among these three meanings, the crucial one for us is the compositional sense because this will provide our explanation for the transitivity restriction on co-fronted subjects.

In this example <u>picture</u> had just one compositional sense, `picture of girl g', because it had just one dependent. What happens if a word has more than one dependent? As in other theories, the final compositional meaning is built in steps rather than in one fell swoop; in dependency terms this means that each dependent defines a different sense. Consider the following pair of examples:
(67)   a.   Pat ate muesli, which Jo never does.
       b.   Pat ate muesli, which usually annoys Jo.
What is the antecedent of <u>which</u>? It must be one of the meanings of the main clause, but which one? In (a) it must be `eating muesli', but in b) it must be `Pat eating muesli' - i.e. combining <u>ate</u> with its object gives one sense, and combining it with its subject gives another. Once again the two meanings are in an isa relationship, with the subject-based sense as a special case of the object-based one. The result is `semantic phrasing' (Hudson 1990:146-51), a hierarchical structure (which may even be binary) in the semantics which maps onto the

completely flat structure in the syntax. The effect is illustrated for <u>Pat ate muesli</u> in **Figure 17**.
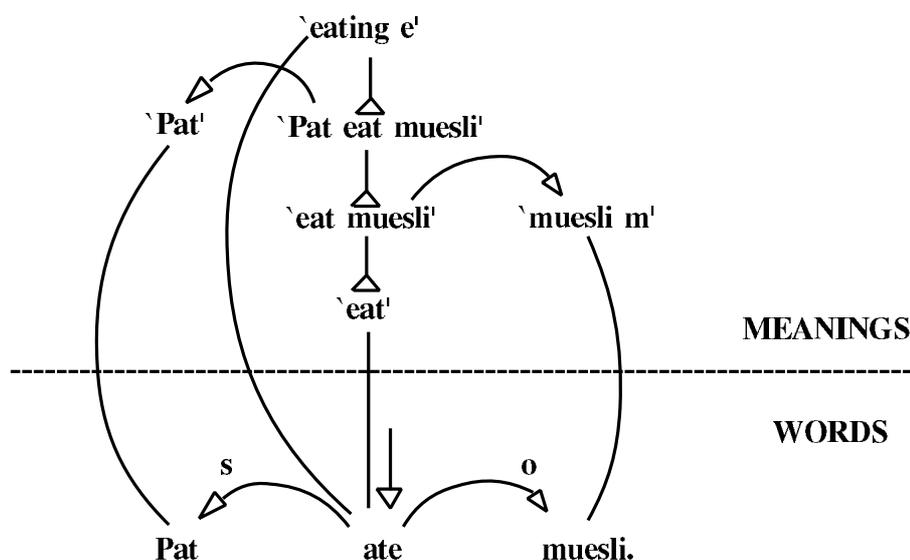

Figure 17

The idea of step-wise composition is important enough to be expressed as a principle:

(68) **The Step-wise Composition Principle**
       Each dependent of a word W may contribute a separate sense of W.

       The most interesting characteristic of these semantic structures is that they reproduce the phrases of PS analysis, including the VP grouping of verb and object. But how is the mapping between flat syntactic structures and hierarchical semantic structures controlled? This is similar to the question about how discontinuous phrases are controlled in the absence of phrase structure, and once again the advantage may at first seem to lie with PS analysis. Just as PS analysis offers a natural explanation for the ban on discontinuity, so it offers a natural foundation for building hierarchical semantic structures: each compositional sense corresponds to a separate syntactic phrase. But this apparent advantage disappears when we consider the flexibility which is available in natural language.

       We shall return shortly to the specially rigid relationship between subjects and objects, but we can first illustrate this flexibility. The perfect illustration is in fact German PVPF, which allows a verb to combine with almost any combination of dependents in order to define a sense. Consider once again Uszkoreit's (1987) virtuoso example from (4).

(69) a.   [Zustecken]$_X$ sollte der Kurier den Brief nachher einem Spion.
             slip should the courier the letter afterwards to.a spy
       b.   [Den Brief zustecken]$_X$ sollte der Kurier nachher einem Spion.
       c.   [Einem Spion zustecken]$_X$ sollte der Kurier nachher den Brief.
       d.   [Nachher einem Spion zustecken]$_X$ sollte der Kurier den Brief.

German allows one to build any of the senses `slip', `slip the letter', `slip to a spy', or `slip afterwards to a spy', as well as all the other mathematically possible combinations of these dependents. This is not at all as predicted by PS analysis.

33

Given the WG approach, on the other hand, it is exactly what one would expect without any further provision. Unless we impose extra restrictions, any combination of dependents may define a compositional sense.

One such restriction applies to subjects and objects. A subject combines with the sense defined by an object, and not vice versa. It is hard[21] to interpret Pat ate muesli in such a way that `Pat eating' is one of its senses which then combines with `muesli'. If objects cannot be combined with subject-based senses, the sense `Pat eating' is not possible. We can call this generalization the Subject-Object Asymmetry Principle.

(70) **The Subject-Object Asymmetry Principle**
A verb's subject may contribute a sense based on the one contributed by its object, but not vice versa.

This principle is the dependency equivalent of the VP node, so it is responsible for many of the subject-object asymmetries which PS analysis explains in terms of syntactic asymmetry. It will also help us to explain the transitivity restrictions on co-fronted subjects in PVPF by explaining the interaction between subjects and objects.

Before we confront these restrictions again we need a little formal apparatus for representing senses. The sense of W on its own is `sense (W)', but compositionality means that this sense is modified by W's dependents, each such modification giving a different sense[22]. For each dependent word $d_i$, we can define a new sense `sense (W | d_i)' (`the sense of W in the context of $d_i$'), and we can generalize across dependents by putting general definitions such as `subject of W' into the formula. Thus the word ate in Pat ate muesli has the three senses in (71), each of which is an example of the more general formula shown below it.

(71) a. sense (eat) = `eat'
            sense (W)
      b. sense (eat | muesli) = `eat muesli'
            sense (W | object (W))
      c. sense (eat | muesli | Pat) = `Pat eat muesli'
            sense (W | object (W) | subject (W))

These three senses form an isa hierarchy:

(72) a. sense (eat | muesli | Pat) isa sense (eat | muesli)
      b. sense (eat | muesli) isa sense (eat)

We can simplify the first formula in (71) by omitting all but the outermost dependent:

(73) sense (eat | Pat) = sense (eat | muesli | Pat)
      i.e. sense (W | subject (W)) = sense (W |object (W) | subject W))
      i.e. sense (W | d_j) = sense (W | d_i | d_j) for each d_i such that sense (W | d_i | d_j) isa sense (W | d_j)

What we now need to explain is why the subject-defined sense has to `isa' the object-defined sense, and how this explains the PVPF facts.

In the last section I introduced the Closing-Subject and Object-first Principles informally, but we can now make them more precise. The Closing-Subject Principle says that the subject `freezes' the verb's sense. In terms of compositional senses, this means that the subject-defined sense must include the senses defined by all the other dependents. (The revised version is in (79).)

(74) The Closing-Subject Principle (first formalized version)

The sense of W defined by its subject is the sense which is defined by all W's other dependents as well;

i.e. sense (W | subject (W)) = sense (W | every dependent (W))

I explained in the last section how this principle predicts that subjects generally freeze the sense of X, so that no further dependent may be added after the finite verb. Take an earlier example repeated here:

(75)   [Ein Außenseiter gewonnen]$_X$ hat da noch nie.

an outsider won has (that) there still never

`An outsider has never yet won there.'

`An outsider won' is fine as a complete sense for gewonnen, but its understood object must have an indefinite or anaphoric interpretation, and this must be the final sense of gewonnen because of the closing effect of its subject `an outsider'. Since the sense of X must already include all the modifications supplied by dependents, it cannot then be changed to accommodate the interpretation required by a later NP such as `the Derby'.

In contrast, non-subjects can combine freely with fronted verbs. For example take the alternatives illustrated in (69). If X is just zustecken, `slip', without any co-fronted dependents, its sense at the point where it counts as the sense of X is just `slip' but there is no presumption that this is the definite sense of zustecken. Consequently the processor can accommodate dependents wherever they occur, whether before or after zustecken.

Why are unaccusative and passive verbs special? My suggestion in section 9 was that it is because their subjects are also their objects, and their object-characteristics take priority. Take ein Fehler unterlaufen, `a mistake happened'. If ein (Fehler) is both the subject and the object of unterlaufen, and if object characteristics win, the Closing-Subject Principle will not apply, so another dependent of unterlaufen may follow it. This needs to be formalised. First, we give the WG version of the Object-First Principle:

(76)   The Object-First Principle (formalized)

The sense of W defined by its object must not include any of W's senses defined by other dependents;

i.e. sense (W | object (W)) ≠ sense (W | other dependent (W))

This leaves a serious problem: how to make sure that this will override the Closing-Subject Principle.

Why should objects take priority over subjects? After all, default inheritance gives priority to A over B provided that A isa B, but nobody would claim that `object' isa `subject'. Worse still, there are other areas of grammar where the relationship is reversed: the characteristics of subjects take priority over those of objects. This is true in matters of case and word order; for example, if we again assume that unaccusative and passive verbs have subjects which are also objects, it is clearly the subject-hood that wins out in terms of case and word order. How can we both reconcile these conflicting priorities, and also explain why there should be any priority at all?

Suppose that subjects and objects are both examples of the same supercategory. This may be the same as the category `argument' or `valent', but for present purposes we don't need to decide this and we can adopt a temporary cover term, `sobject'. (Though this may include other kinds of argument, it certainly excludes adjuncts and predicatives (our `sharers').) All we need, therefore, is the
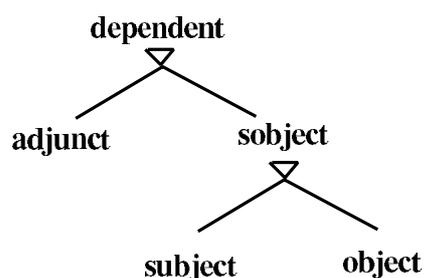
35

analysis in **Figure 18**.

Figure 18

This allows an easy explanation for why subject characteristics sometimes take priority over object ones, while on other occastions the reverse is true. Take case, for example. An object is accusative unless it is also a subject (as with unaccusative and passive verbs), and subjects are always nominative. In terms of default inheritance, accusative is the default which is overridden by nominative, so we formulate the case rules so that sobjects (not objects) have accusative case and subjects nominative.

(77)   Case
       a. Sobjects are accusative.
       b. Subjects are nominative.

By way of contrast, linking rules tend to work the other way round, with object taking priority; so passive subjects, which are also objects, are semantically like objects rather than subjects. For simplicity we can assume that there are just two semantic roles, agent and patient.

(78)   Linking
       a. Sobjects are agents.
       b. Objects are patients.

These rules take `agent' as the default for sobjects, which is overridden if the sobject is more precisely an object.

We now have the apparatus that we need for an explanation of the restrictions on co-fronted subjects. The general question was why unaccusative and passive subjects don't freeze the sense of the fronted verb in the way that other subjects do; more specifically, why does the Object-First Principle take priority over the Closing-Subject Principle? The answer is that the rule for subjects provides the default, so this must in fact apply to `sobject', not to `subject'; i.e.

(79)   The Closing-Subject Principle (final formalized version)
       The sense of W defined by its sobject must include W's senses defined by
       all its other dependents;
          i.e. sense (W | sobject (W)) = sense (W | every dependent (W))

The revised Closing-Subject Principle applies to sobjects in general but will always be overridden for objects by the Object-First Principle, regardless of whether or not they are also subjects.

To summarise this section, I have provided a semi-formalised WG analysis for co-fronted subjects which explains the special freedom of passive and unaccasative verbs to have later, un-fronted, dependents. The Step-wise Combination Principle allows fronted verbs to combine freely with any of their co-

fronted dependents to produce a distinct `sense' in the semantic structure. This sense is the clause's topic because it is the meaning of X at the point where it is first processed, but it may be supplemented later by further senses triggered by further dependents. However this freedom is restricted by the Closing-Subject Principle because a co-fronted subject (more precisely, a `sobject') blocks the addition of later senses for X; but this principle is in turn overridden by the Object-First Principle if the subject happens also to be an object.

11.   Conclusion

The WG rules and principles have solved most of the problems of German Partial VP Fronting. They allow PVPF in general by allowing generalized raising, a rule for German which allows any dependent of a fronted non-finite verb to depend on the higher verb as well as on the lower one. They allow subjects to be co-fronted with the non-finite verb by a special relaxation of the Raising Principle, which again is a peculiarity of German grammar. And they restrict this co-fronting of subjects by means of principles and rules that control the interface between syntax and semantics. Admittedly there are some loose ends (the exact formulation of generalized raising, the treatment of pronoun objects after the finite verb and the definiteness restriction on co-fronted subjects) but the analysis has gone beyond any earlier analysis of PVPF.

So what? PVPF is just a tiny corner of the grammar of one language. Even if we include other languages that have PVPF (Dutch and maybe others), the achievement is quite limited. Nevertheless it is possible to draw a conclusion of some importance for the general theory of grammar: to the extent that this analysis has been more successful than earlier attempts it is at least partly because I have used dependency structure rather than phrase structure. This has provided just the right degree of flexibility needed to accommodate PVPF, without opening the floodgates of overgeneration. I think it is fair to say that this combination of general principles and learnable parochial rules has explained PVPF in a way that has not yet been possible with phrase structure. The balance of explanatory power may not in fact be as seen by Pollard and Sag (1994:10):

.. for all that a theory that successfully dispenses with a notion of surface constituent structure is to be preferred (other things being equal, of course), the explanatory power of such a notion is too great for many syntacticians to be willing to relinquish it.

Department of Phonetics and Linguistics,
University College London,
Gower Street,
London WC1E 6BT
dick@ling.ucl.ac.uk

References

Baker, Kathryn. 1994. An extended account of `Modal Flip' and Partial Verb Phrase Fronting in German. Carnegie Mellon University Laboratory for Computational Linguistics: Technical Report CMU-LCL-94-4.

Bech, Gunnar. 1955. Studien über das deutsche verbum infinitivum, vol 1, 2nd edition. Tübingen: Niemeyer.

den Besten, Hans and Webelhuth, Gert. 1990. Stranding. In Grewendorf and Sternefeld (1990:77-92).

Chomsky, Noam. 1981. Lectures on Government and Binding. Dordrecht: Foris.

Chomsky, Noam. 1986. Barriers. Cambridge, Mass: MIT Press.

Chomsky, Noam. 1995. The Minimalist Program. Cambridge, MA: MIT Press.

Durrell, Martin. 1996. Hammer's German Grammar and Usage, 3rd edition. London: Arnold.

Fraser, Norman. 1994. Dependency Grammar. In Encyclopedia of Language and Linguistics, ed. Ronald Asher, 860-4. Oxford: Pergamon.

Gazdar, Gerald; Klein, Ewan; Pullum, Geoffrey and Sag, Ivan. 1985. Generalized Phrase Structure Grammar. Oxford: Blackwell.

Grewendorf, Günther. 1988. Aspekte der Deutschen Syntax. Tübingen: Gunter Narr.

Grewendorf, Günther and Sternefeld, Wolfgang. eds. 1990. Scrambling and Barriers. Amsterdam: Benjamins.

Haegeman, Liliane. 1991. Introduction to Government and Binding Theory. Oxford: Blackwell.

Haegeman, Liliane. 1991/94. Introduction to Government and Binding Theory. Second Edition. Oxford: Blackwell.

Haider, Hubert. 1981. Empty categories: on some differences between German and English. Wiener Linguistische Gazette 25:13-36.

Haider, Hubert. 1990. Topicalization and other puzzles of German syntax. In Scrambling and Barriers, eds. Gunther Grewendorf and Wolfgang Sternefeld, 93-112. Amsterdam: Benjamins.

Haider, Hubert. 1993. Deutsche Syntax - generativ. Vorstudien zur Theorie einer projectiven Grammatik. Tübingen: Gunter Narr.

Haider, Hubert; Olsen, Susan and Vikner, Sten. eds. 1995. Studies in Comparative Germanic Syntax. Dordrecht: Kluwer.

Hinrichs, Erhard and Nakazawa, Tsumako. 1994. Partial VP and split NP topicalization in German: An HPSG analysis. Unpublished manuscript.

Hopper, Paul and Thompson, Sandra. 1980. Transitivity in grammar and discourse. Language 56:251-99.

Hudson, Richard. 1984. Word Grammar. Oxford: Blackwell.

Hudson, Richard. 1990. English Word Grammar. Oxford: Blackwell.

Hudson, Richard. 1993. Do we have heads in our minds? In Heads in Grammatical Theory, eds. Greville Corbett, Scot McGlashan and Norman Fraser, 266-91. Cambridge: Cambridge University Press.

Hudson, Richard. 1994a. Word Grammar. In The Encyclopedia of Language and Linguistics, ed. Ronald Asher, 4990-3. Oxford: Pergamon.

Hudson, Richard. 1994b. Discontinuous phrases in dependency grammar. UCL Working Papers in Linguistics 6:89-124.

Hudson, Richard. 1995a. <u>Word Meaning</u>. London: Routledge.

Hudson, Richard. 1995b. Really bare phrase-structure = dependency structure. <u>Studies in English Usage and English Language Teaching 17:</u>3-17.

Hudson, Richard. 1997a. The rise of auxiliary DO: verb-raising or category-strengthening? <u>Transactions of the Philological Society 95:</u>41-72.

Hudson, Richard. 1997b. Inherent variability and linguistic theory. <u>Cognitive Linguistics 8:</u>73-108.

Hudson, Richard. 1997c. Syntax without functional categories. <u>UCL Working Papers in Linguistics 9:</u>

Hudson, Richard. forthcoming. <u>English Grammar</u>. London: Routledge.

Johnson, Mark. 1986. A GPSG account of VP structure in German. <u>Linguistics 24:</u>871-882.

Judge, Ann and Healey, F. G. 1985. <u>A Reference Grammar of Modern French</u>. London: Arnold.

Kayne, Richard and Pollock, Jean-Yves. 1978. Stylistic inversion, successive cyclicity and move-NP in French. <u>Linguistic Inquiry 9:</u>595-621.

Keenan, Edward and Comrie, Bernard. 1977. Noun-phrase accessibility and universal grammar. <u>Linguistic Inquiry 8:</u>63-99.

Müller, Stefan. 1995. Yet another paper about Partial Verb Phrase Fronting in German. Unpublished paper.

Nerbonne, John. 1986. `Phantoms' and German fronting: poltergeist constituents? <u>Linguistics 24:</u>857-70.

Nerbonne, John. 1994. Partial verb phrases and spurious ambiguities. In <u>German in Head-driven Phrase Structure Grammar (CSLI Lecture Notes 46)</u>, eds. John Nerbonne, Klaus Netter and Carl Pollard, 109-50. Stanford: CSLI.

Netter, Klaus. 1991. <u>Clause union and verb raising phenomena in German (Research Report 91-21)</u>. Saarbrücken: Deutsches Forschungszentrum für Künstliche Intelligenz.

Pollard, Carl and Sag, Ivan. 1994. <u>Head-driven Phrase Structure Grammar</u>. Chicago: University of Chicago Press.

Prinzhorn, Martin. 1990. Head movement and scrambling domains. In Grewendorf and Sternefeld (1990:199-215).

Radford, Andrew. 1997. <u>Syntactic Theory and the Structure of English. A minimalist approach</u>. Cambridge: Cambridge University Press.

Roberts, Ian. 1993. <u>Verbs and Diachronic Syntax. A comparative history of English and French</u>. Dordrecht: Kluwer.

Rosta, Andrew. 1994. Dependency and grammatical relations. <u>UCL Working Papers in Linguistics 6:</u>219-58.

Sigurðsson, Halldór. 1991. Icelandic Case-marked PRO and the licensing of case-marked arguments. <u>Natural Language and Linguistic Theory 9:</u>327-64.

Tesnière, Lucien. 1959/65. <u>Éléments de Syntaxe Structurale.</u> Paris: Klincksieck.

Tomlin, Russell. 1986. <u>Basic Word Order: Functional Principles</u>. London: Croom Helm.

Uszkoreit, Hans. 1987. Linear precedence in discontinuous constituents: complex fronting in German. In <u>Discontinuous Constituents (Syntax and Semantics 20)</u>, ed. Geoffrey Huck and Almerindo Ojeda, 405-25. San Diego: Academic Press.

Webelhuth, Gert. 1990. Diagnostics for structure. In Grewendorf and Sternefeld (1990:41-75).

1. The problem addressed in this paper has been on my mind since it was first pointed out to me by Martin Prior about twenty years ago. It resurfaced ten years ago when I read Hans Uszkoreit's seminal paper (1987), and I have been thinking about it on and off since then. The solution that I present here has developed through an unpublished paper `HPSG without PS?' (1995) and a presentation to the Linguistics Association of Great Britain in September 1996. A lot of people have contributed to it in one way or another, but I should like to pick out Avery Andrews, Kathryn Baker, Sylvain Kahane, Andreas Kathol, Klaus Netter and Stefan Müller for special thanks.

2. Default inheritance should not be confused with the Chomskian notion of `inheriting' features or barrier-properties (Haegeman 1991/94:556). By Chomskian inheritance a mother node in a sentence structure receives a property from its complement; for example, DP may inherit features from its complement NP, and CP from its IP. In contrast, default inheritance is a relationship between two nodes in an `isa' or `generalization' hierarchy whereby the more specific node inherits all the characteristics of the more general one by default, i.e. in the absence of any more specific characteristics. The notion is familiar in Artificial Intelligence as well as in some current linguistic theories (Pollard and Sag 1994:36).

3. This is a slight oversimplification, as other constraints have been suggested. For example, Grewendorf (1988:297) gives the following judgements which do not seem to be taken seriously by other linguists:
(i)   a   [Auf den Tisch gestellt] hat Maria die Vase.
                on the table put has Mary the vase.
                `Mary has put the vase on the table.'
      b   *[Die Vase gestellt] hat Maria auf den Tisch.
(ii)   a   [Ein dickes Buch schreiben] wollte Gisbert in Passau.
                a thick book write wanted Gisbert in Passau
                `Gisbert wanted to write a thick book in Passau.'
      b   *[In Passau schreiben] wollte Gisbert ein dickes Buch.
I have also ignored the rather similar claim in Haider (1993:154) about object pronouns, to the effect that an accusative pronoun must give way to a dative one in fronting (though the same is not true of full NPs):
(iii)   a   daß er $sich_{Acc}$ ihr_{Dat} vorgestellt hat
                that he himself to.her introduced had
      b   *daß er ihr_{Dat} $sich_{Acc}$ vorgestellt hat
      c   [Ihr_{Dat} vorgestellt] hat er sich nicht.
      d   *[$Sich_{Acc}$ vorgestellt] hat er ihr nicht.
These examples clearly deserve more attention.

4. The terms `coherent' and `incoherent' are explained in Grewendorf and Sternefeld (1990:9), who attribute them to Bech (1955). A construction in which one verb depends on another is coherent if the two verbs are clause-mates, and incoherent if they are not. In the terminology that I shall use in later sections, only coherent constructions involve `sharers'.

5. Thanks to Andreas Kathol, Stefan Müller and Tibor Kiss, who all independently confirmed that the finite verb agrees with the co-fronted subject.

6. Both these examples were provided by Stefan Müller.

7. The data are somewhat unclear in this area. Webelhuth (1990:53) and Grewendorf (1988:298) are uncertain about the grammaticality of the examples in (i), although Haider rejects such examples out of hand.
(i)   a   ?[Ein alter Mann gelesen] hat das Buch niemals.
                an old man read has the book never
                `An old man has never read the book.'
         b   ?[Deutsche gewonnen] haben Wimbledon noch nie.
                Germans won have Wimbledon still never
                `Germans have not yet won Wimbledon.'
The two important features of both these examples are that the non-fronted object is a full NP and that it is immediately next to the finite verb. Furthermore Haider (1993:153) gives a `?' to (ii), in which `eaten' seems to have quite high transitivity (though the generic `lobsters' is not a prototypical object). The WG analysis offered later actually predicts that examples like this should be acceptable, though they conflict with the preliminary generalisation in terms of transitivity.
(ii)   [Linguisten Langusten gespeist] haben da wohl noch nie.
         linguists lobsters eaten have there presumably still never
         `Linguists have presumably not yet eaten lobsters there.'

8. `Pronouns normally follow immediately after the finite verb or the conjunction' (Durrell 1996:462). More precisely, according to Haider et al (1995:9) pronouns which depend on a coherent complement verb may take their position from the higher verb. Their examples are the following.
(i)   a.   daß sie nicht wagt, ihn zu stören
                that she not dares him to disturb
                `that she does not dare to disturb him'
         b.   daß sie ihn nicht zu stören wagt
In example (b), `not' depends on `dares' but separates `him' from its parent verb `to disturb'; this indicates that `not' and `him' are co-dependents of `dares'.

9. Haider (1993:69) raises the same question in relation to PVPF, but does not apply it to scrambling.

10. `Generalized raising' is a better name than `clause union', because the union need not be complete - part of the fronted clause may be united with the main clause while another part is exclusive to X; and `argument attraction' fails because it excludes adjuncts.

11. The term `parent' is better than `head', which I used to use but which is very confusing for those with a background in phrase structure. The head of a phrase is the parent of each of the words that depends on it.

12. The abbreviations are those that I use in my elementary teaching and in a textbook (Hudson forthcoming).

13. I have argued elsewhere (especially in Hudson 1997c) that the `complementizer' <u>that</u> is in fact sui generis, a syncategorematic word.

14. The V2 requirement is actually too strict because some sentence adjuncts such as `in other words' or `after all' don't count (Durrell 1996:455). The WG rule restricts the number of X's to one, so these extra adjuncts can be allowed so long as they are not labelled `X'.

15. In Hudson 1990:43, I argued that a default proposition should apply except when explicitly contradicted by a special proposition `not (p)'. I now believe this was wrong, so I agree with other users of default inheritance, in assuming `automatic overriding': roughly speaking, any default proposition is automatically overridden by a proposition which refers to a more specific concept.

16. I follow the judgements of Sylvain Kahane, who also persuaded me that the examples did indeed involve lowering. However there seems to be a data problem according to Judge and Healey (1985:350):
> The following is probably unacceptable to some:
>> ?La maison qu'a achetée mon père à son voisin est très grande.
> The reason the latter is not totally acceptable is because there is a rule in French which states that one should not place side by side (or one after the other) two phrases which are not grammatically linked: since <u>mon père</u> and <u>à son voisin</u> are not grammatically dependent one on the other, they should not be contiguous. But such constructions do occur in spontaneous spoken French.

17. Something like the Closing-subject Principle is hinted at by Haider (1990:106):
> The verbal projection that contains the non-ergative subject is closed for the projection principle. What [this] amounts to is that a θ-role cannot be passed on to an argument outside a VP that contains a non-ergative subject. Hence in (i) the object remains θ-less and violates the θ-criterion.
> (i) *Ein Außenseiter gewonnen hat da noch nie das Derby.
>> (an outsider won has there still never the Derby)

18. Haider (1990:110) suggests an explanation for the difference between German and Dutch: the subject is internal to VP in German, but not in Dutch. This explanation may need to be rethought in the light of the VP-internal subject hypothesis, which seems to be widely accepted among Chomskian linguists as part of Universal Grammar. (For a review of the evidence, see Radford 1997:318-24.)

19. The straight and curved lines are iconic for the relationships concerned. We can go straight from a word to its sense by simply looking it up in our mental lexicon, whereas we approach its referent only indirectly via pragmatic inference.

20. Again the notation is iconic. The base of the triangle is larger than its apex in the same way that the supercategory (on whose name the base rests) is larger than the individual instances (to which the apex is linked).

21. Though hard it is not impossible to imagine a situation where `Pat eat' would be a useful meaning - sentence (i), for example.
(i) Pat ate muesli, which also happened to toast.
But this is very odd, and hardly provides clear evidence against the constraint.

22. In Hudson 1990:148-51 I suggested a new relationship `head-sense' between a dependent and the sense of its parent (`head' in 1990 terminology). This suggestion was unsatisfactory because it suggested that the parent's sense actually belonged to the dependent. The present suggestion is meant to replace the idea of head-senses.