

Notes on the ECP in English and German

Josef Bayer

Max-Planck-Institut für Psycholinguistik, Nijmegen

October 1989

*For discussions about and/or comments on an earlier version of this paper I am indebted to Lyn Frazier, Hans de Koot, Angelika Kratzer, Gereon Müller, Janet Randall, Elisabeth Selkirk and Wolfgang Sternefeld. None of them can be held responsible for the misconceptions which may appear in the text. Thanks to Yves Fuchs for his help in preparing a readable version.

0. Introduction

My goal in this article is to trace certain differences between English and German (and Dutch) which arise in connection with the Empty Category Principle (ECP). Most of these differences have been known to the community of syntacticians for quite a while, but only recently has the context shifted to allow them to spur a theoretically fruitful exchange. This is a result of the gradual shift in the analysis of English towards a phrase structure resembling, in much detail, the phrase structure assumed for German. While in the earlier days of generative grammar idiosyncratic rules such as 'subject-AUX-inversion' and 'do-support' were taken to account for the facts, these phenomena are now being studied in a framework which allows us to make challenging cross-linguistic comparisons. Analyzing WH-movement and the system of complementation along the lines proposed for German by Thiersch (1978) has led to a considerable sophistication. In the 'Barriers'-framework, Chomsky (1986) assumes an elaborated version of precisely the mechanism proposed by Thiersch, who had reinterpreted the so-called 'topological fields' theory of traditional German syntax in generative transformational terms. Chomsky's further step in 1986 was to make functional categories such as I (=INFL) and C (=COMP) into heads in the sense of a generalized X'-schema. "COMP" is then not a single clause-initial constituent, as in the earlier work of Bresnan, but an articulated configuration with a head and a specifier. Assuming such a structure opens up the possibility of analyzing WH-movement in a novel way, which Chomsky (1986) makes crucial use of. On the other hand, it appears that the generalized X'-schema leads to new problems which have to be overcome, if we are to maintain the overall elegance and explanatory depth of the theory.

The article is organized as follows: in 1. some elementary notions of the Barriers-theory will be introduced. In 2. we will point to a difficulty in explaining the **that**-trace effect in English, as compared to German. The first major section is 3. where the 'Vacuous Movement Hypothesis' is suggested to derive essential differences between English and German without the involvement of the 'Minimality Condition'. The second major section, 4., discusses WH-islands in both languages. Throughout this section it will be demonstrated that lexical government is not sufficient to automatically void ECP-effects in German. In section 5 we will take up this issue in a somewhat speculative way. The conclusion contains a suggestion for the reformulation of the ECP which can account for the differences and similarities discussed. In an appendix, a comparison is made between the approach presented here and recent work by Guglielmo Cinque and Luigi Rizzi.

1. The 'Barriers' Framework

The Barriers theory will be introduced only to a degree necessary to make the argumentation in this article comprehensible. For a deeper understanding it is, of course, necessary to consult Chomsky's own (1986) text. Theoretical assumptions such as 'adjunction to VP in order to dissolve a barrier' will be assumed for the purpose of the argumentation, but it

should be clear that this goal could also be implemented in terms of related approaches such as Kayne's (1983) theory of connectedness or Koster's (1986) theory of 'domain extension'. I will use the Barriers framework here in order to make my claims about the ECP immediately contrastable with the work of Chomsky and Rizzi.

The ECP as proposed in Chomsky (1981) and related work involves two conditions which may apply disjunctively. One condition says that an empty category is licit when it is governed by a lexical category; the other says that it is licit when it is governed by a local antecedent. Let us state the ECP in a more recent version which replaces 'lexical government' by the notion 'theta government':

(1) The Empty Category Principle (ECP)

α properly governs β iff α θ -governs or antecedent-governs β .

(Chomsky, 1986:17)

In Chomsky (1986), ' θ -government' is defined as follows:

(2) θ -government:

α θ -governs β iff α is a zero-level category that θ -marks β and α and β are sisters

'Antecedent government' in the Barriers-framework is constrained by the fact that maximal categories such as NP, VP, PP etc. are blocking categories and act as barriers to the government of a trace from outside unless they can be dissolved in one way or the other. We do not want to outline the Barriers theory here (as stated in Chomsky, 1986), but just repeat the central definitions, namely those of the notion 'Blocking Category' (BC) and 'Barrier':

(3) Blocking Category (BC)

γ is a BC for β iff γ is not L-marked and γ dominates β .

(4) Barrier

γ is a barrier for β iff (a) or (b):

a. γ immediately dominates δ , δ a BC for β .

b. γ is a BC for β , $\gamma \neq \text{IP}$.

'Domination' by a node β is taken to hold for some category α only if α is dominated by every segment of β . This requirement yields the result that in an adjoined structure such as $[_\beta \alpha [_\beta \dots]]$, α will not count as being dominated by β . By these prerequisites it is guaranteed that once adjunction to a maximal category β can take place (where β is not L-marked), β will cease to be a BC and thus not count as a barrier. What this means for antecedent government is that for the core cases a tightly linked chain can be built through which an empty category (which may not be θ -governed) can be reached in its minimal maximal category.

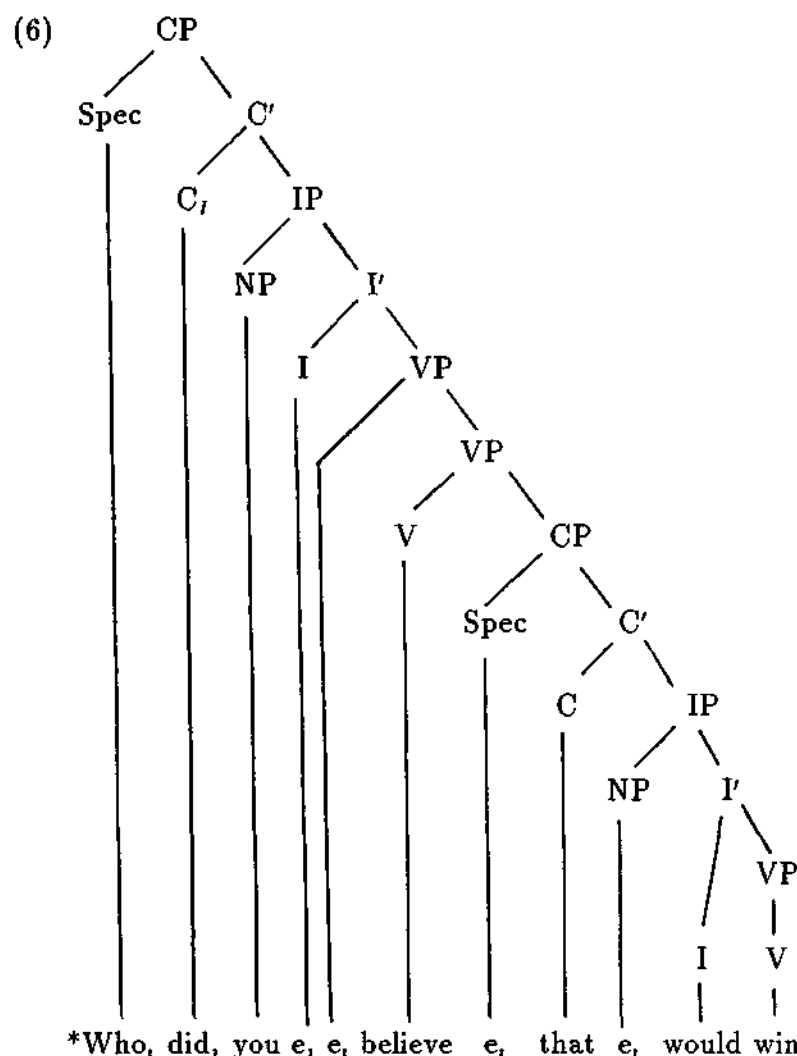
2. that-trace

The phrase structure employed in the Barriers system is such that the functional categories I(NFL) and C(OMP) project into head, complement and specifier. Thus, that heads

a CP of which an IP is the complement and SpecCP the specifier. This specifier can be used as an “escape hatch”, to use more traditional terminology. A trace inside IP that is not θ -governed can still circumvent the ECP when there is an intermediate trace in SpecCP which can serve as an antecedent governor. This predicts the well-formedness of cases such as (5), all irrelevant details aside.

(5) How, do you think [e, [that [John fixed his car e,]]]?

This creates a problem, however, because a chain-explanation of the **that**-trace effect will be lost. Consider a classical case of a **that**-trace violation and the structure given in (6).



According to the assumptions in Barriers, IP is not a barrier. So the subject-trace under the lower IP can be governed from outside and movement from this position can proceed. The intermediate trace in the lower specifier position is also allowed, since it remains within a minimal maximal projection, namely CP.¹ The CP in turn, by being L-marked (by the verb **believe**) is not a BC and thus not a barrier (see (4b) above). The higher IP cannot be a BC, because it is L-marked by its own (moved) head. Thus, movement can proceed, and a

standard example of grammatical long WH-extraction is predicted: An undesirable result, because (6) shows a *that*-trace violation. Something else must be said to rule such cases out.

The conditions posited so far have been identical for both movement and government. One could try to rule out cases such as (6) by proposing a stricter version of government. This is exactly Chomsky's tack in chapter 8 of *Barriers*, where the 'Minimality Condition' (MC) is invoked in order to account for the *that*-trace effect. Without trying to go into the various ramifications of the MC, let us just assume the following simplified (strong) version which is supposed to hold for government, but not for movement:

(7) **The Minimality Condition**

In a structure ... β ... [δ γ ... α ...], where δ is a projection of γ that dominates γ immediately, α cannot be governed by β .

'Projection' in (7) must be understood here as any projection. According to (7), (6) will be ruled out as an ECP-violation in the following way: The lower category C' , which is headed by *that*, immediately dominates the category in which the subject-trace e_i is found. Thus, antecedent government by the intermediate e_i is blocked by the complementizer, which may be taken to be something like a "closer governor".² Chomsky assumes that only interveners with features (i.e. γ in (7)) can count as minimality barriers (see *Barriers*, pp.47f). Phonetically empty heads such as zero-complementizers are taken to be featureless. As a consequence, when *that* is absent, the MC is not violated, and antecedent government of the subject-trace will be possible.

Chomsky himself points out a problem for this approach with respect to the cyclic movement of adjuncts: Unless I' (besides IP) is taken to be a defective category, I could be taken to induce a minimality barrier for the lower adjunct-trace in (8) (= (103) in *Barriers*):

(8) [$_{CP}$ e' [$_{IP}$ NP [$_{I'}$ I [$_{VP}$ e [$_{IP}$...]]]]]

There is another problem also: Consider the case of adjunct movement in (5) again. In the same way as the MC can be invoked to derive an ECP-effect for (6), it can be invoked to derive one for (5). Since the trace of *how* is not lexically governed, antecedent government should not be blocked by an intervening complementizer. But even in the presence of the complementizer, (5) is grammatical. This problem was independently noticed in Rizzi (1987).³

A third problem arises from the syntax of German and Dutch. In both languages the *that*-trace effect appears to be absent. On the other hand, there is clear evidence that the projection of an embedded clause is indeed as assumed in the *Barriers* framework. Certain Dutch and German dialects exhibit overt movement of a WH-element to the SpecCP-position when the complementizer *dat*/*daß* is present in C-position. Consider, for instance the English/South-German contrast in (9).

- (9) a. *John does not know [who, [that [e_i , will win]]]
 b. Hans weiß nicht [wer, [daß [e_i , gewinnen wird]]]

Why should there be no **that**-trace effect in these languages? The MC as stated above would predict one.

Rizzi (1987) does not share Chomsky's proposal that **that** sets up a minimality barrier. Instead, Rizzi suggests that the **that**-trace effect arises in English because the subject-NP is not **canonically** governed by I, while it is canonically governed by I in Dutch and German. We will show below that this explanation faces problems, too. We agree with Rizzi that the MC does not seem to be the appropriate generalization for capturing the **that**-trace effect. Let us therefore pursue a different line of theorizing.

3. Vacuous Movement

In this section we want to propose that the so-called 'Vacuous Movement Hypothesis' (VMH), originally proposed by Chomsky (1973) and George (1980) and taken up again by Chomsky (1986), can serve to derive an important difference between English on the one hand and German/Dutch on the other.

According to Chomsky (1986), the VMH says that a transformation that applies string-vacuously may be treated as not applying at all. The situation is exemplified in (10).

(10)a. Non-vacuous movement

(i)...Z...X...Y... \Rightarrow (ii)...X...Z...Y... , i.e. (ii) \neq (i)

b. Vacuous movement

(i)... \emptyset ...X...Y... \Rightarrow (ii)...X... \emptyset ...Y... , i.e. (ii)=(i)

In (10a) X is visibly/audibly disconnected from Y by the intervening factor Z. In (10b), however, the factor which splits X and Y is phonetically empty. In the cases we will consider it is a constituent boundary. For the child who is exposed to data that may involve a string-vacuous rule application, the rule appears as not having applied at all. There is simply no positive evidence that it has.

Examples provided by George include cases of coordination, i.e. structures which require strict parallelism. Notice the following contrast:

(11)a. They removed the prisoner, whom the judge has sentenced and whom the warden will execute

b. *They removed the prisoner, who has lost his appeal and whom the warden will execute

The VMH predicts that the coordinated relatives in (11a) are parallel, but those in (11b) are not. The reason is that the subject **who** in contrast to the object **whom** is moved only string-vacuously. Under the assumption that there is no string-vacuous movement, the output of the two WH-movements in (11b) is given in (11'):

(11')a. [_{IP} who has lost his appeal]

b. [_{CP} whom_i [_{C'} [_{IP} the warden ...e_i]]]

According to standard assumptions about coordination the two clauses of (11') cannot be conjoined because they are of different categorial status. In this way, the VMH is able to predict the ill-formedness of (11b) in contrast to the well-formedness of (11a).

While Chomsky (1986) assumes that "vacuous movement is not obligatory at S-structure" (pp. 49f.), George (1980) adopts it as obligatory.⁴ It is hard to see how under the given assumptions the contrast in (11) could be derived if the VMH is not taken to be obligatory. There would always be a derivation at hand which makes (11b) grammatical. So let us stick to George's original proposal that movement **must not** apply (at S-structure) in case the language never exhibits string-changing operations in a particular construction type. The qualification "at S-structure" is important, because as Chomsky (1986) points out, in order to derive interpretable LFs, movement must still occur at LF. Consider his pair of examples in (12) below.

- (12)a. Who (___) likes John? (vacuous)
 b. Who does John like ___? (non-vacuous)

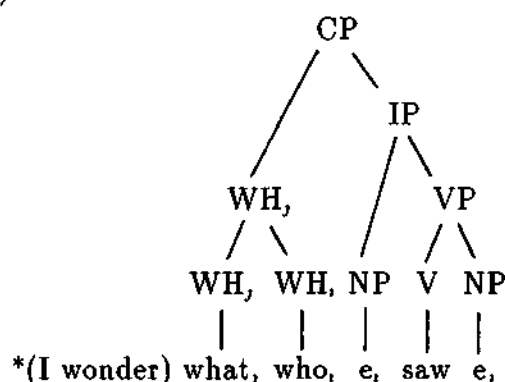
In (12b) movement must be assumed because the WH-object is stringwise displaced from its D-structure position. (12a) instead is consistent with an in-situ analysis of the WH-word. Chomsky assumes that in a case such as (12a) movement to SpecCP will only take place at the level of LF, thus satisfying the operator/variable requirement for WH-interrogatives.

Chomsky demonstrates this, among other examples, with the ill-formedness of examples such as (13) (= (108) in *Barriers*):

- (13) *I wonder [_{CP} what [_{C'} [_{IP} who saw]]]

Once **what** occupies SpecCP, it is not possible for **who** to move to this position in LF. Following May (1985) one can assume that in such a case of multiple interrogation the LF-moved WH-item, in (13) **who**, adjoins to the WH-item which was moved in the syntax proper. This gives an LF-representation as follows in (14).

(14)



(14) shows a superiority violation which either reduces to an ECP-violation (as Chomsky argues) or to a violation of Pesetsky's (1982) "Path Containment Condition" (as May argues).⁵

One problem for VM pointed out by Gereon Müller and Wolfgang Sternefeld (p.c.) is that in **that**-relatives it seems unclear what the nature of the subject should be, if it is not a

trace. One way of overcoming this problem within a VM-approach is based on the fact that the complementizer in cases like **the thing that troubles me** is formally identical with the pronominal **that**. At S-structure the subject could be realized by this pronominal, while at LF **that** is moved to C-position taking on its property as a complementizer which contains an empty operator. This operator moves to SpecCP and binds the variable in subject-position. To do so, **subject-that**, **complementizer-that** and empty operator must share one index. Under these assumptions the VMH in its strong interpretation remains intact.

There are other questions concerning the VMH, some of which were pointed out by Clements, McCloskey, Maling and Zaenen (1983). In the absence of further evidence about VM in the languages studied by these authors, I can only accept their criticism.

We will now turn to German, and argue that this language is unlikely to give rise to string-vacuous derivations of the sort we have found in English. German does not make a distinction between subjects and objects in forming WH-interrogatives, as shown by (15).

- (15)a. Wer, hat e, den Hans gesehen? (non-vacuous)
 Who has the H. seen
 "Who has seen Hans?"
 b. Wen, hat der Hans e, gesehen? (non-vacuous)
 "Who(m) has Hans seen?"

Given that the D-structure is [SUBJECT] [OBJECT] **gesehen hat**, and a Verb Second Rule (V2) moves the finite verb to the C-position, both object- (see 15b) and subject-NP (see 15a) are non-vacuously moved in the syntax. Both appear in SpecCP. In short, movement of the WH-element is never string-vacuous, and the examples appear to be irrelevant for the VMH. One could, however, argue that the VMH may still be relevant for German embedded WH-clauses. Notice that if (15a) is embedded, it will have a V-final structure, and thus the finite verb does not interrupt the sequence COMP + SUBJECT. This is shown in (16a). (16b) shows the same thing with respect to the formation of a relative clause:

- (16)a. Arabella weiß [wer den Hans gesehen hat]
 Arabella knows who Hans seen has
 "Arabella knows who has seen Hans"
 b. die Frau [die den Hans gesehen hat]
 the woman who Hans seen has
 "the woman who has seen Hans"

The child acquiring German could assume that on the basis of such data there is no movement. This would make data like those in (16) consistent with our strong interpretation of the VMH. Notice, however, that many varieties of spoken German allow for structures such as (9b) above or (17) below:

- (17)a. Arabella weiß [wer [daß [den Hans gesehen hat]]]
 b. die Frau [die [wo [den Hans gesehen hat]]]

In both of these cases the subject has been moved across the complementizer, in (17a) across *daß*, in (17b) across the complementizer *wo*, which is used for relativization. In Dutch, WH-complements are often introduced by the complementizer *of* ('whether') or even by *of dat* ('whether that') which are then preceded by the WH-subject. Koster (1986) provides the following example:

- (18) Ik weet niet wie, of (dat) e, het gedaan heeft
 I know not who whether(that) it done has
 "I don't know who did it"

We cannot, of course, exclude that on the basis of data such as (16) and parallel data in Dutch the VMH becomes effective in Dutch and German as well. Empirical considerations, however, suggest that this is unlikely, and that it is ineffective in the mature competence system. This is shown by the acceptability of a translation of (11b):

- (19) Sie entfernten den Gefangenen, der verurteilt wurde und den der Aufseher
 They removed the prisoner who sentenced was and whom the warden
 hinrichten wird
 execute will

The nominative d-pronoun *der* seems to be in a position parallel to the dative d-pronoun *den*. Similar contrasts between English and German can be observed in examples that are discussed in Chung and McCloskey (1983) for English. Chung and McCloskey argue in favor of a GPSG account of relativization according to which a slash category is associated with objects but not with subjects. Consider their examples in (18) for which we give somewhat simplified representations:

- (20)a. That's one trick that, I've known [a lot of people who've been taken in by e,
 b. *That's one trick that, I know [a lot of people that, the police have taken in e,
 with e.]

In (20b) the Wh-item associated with the trace and the complementizer relativizing on *trick* has to cross two bounding nodes, NP and CP, yielding a subjacency violation. Assuming VM in (20a), however, the relativization operating on *people* keeps SpecCP free, thus permitting cyclic movement of the WH-item to take place. In this case only one bounding node will be crossed, and the structure is felt to be not or at least much less offending. Looking at similar examples in colloquial northern and western German, one does not find such a contrast. Irrespective of the nature of the argument involved in relativization they are ungrammatical.

- (21)a. *Da, kenne ich [viele Leute die [_{PP} e, von] begeistert sind
 that know I many people who of enthusiastic are
 "I know many people who(nominative) are fond of that"
 b. *Da, kenne ich [viele Leute denen, e, [_{PP} e, von] schlecht wird
 that know I many people to-whom of sick become

"I know many people who(dative) get sick by that"

Independent evidence for the non-applicability of the VMH in Dutch is provided by van de Koot (1988), who contrasts superiority effects and WH-island violations in English and Dutch. We will turn to WH-island violations below. For the time being it seems safe to assume that at least for subject-WH-movement the VMH is not applicable in Dutch and German while there are good reasons in favor of it in English.

Let us now turn back to the **that**-trace effect. By the VMH there is no reason why in English WH-movement of a subject will ever access SpecCP. It is more likely that the CP prunes down to C', thus containing the complementizer **that** but nothing like a specifier position. We will make this more precise below. At the moment it should be noted that if the specifier position were to stay, it would have to be constrained in some way; otherwise material from inside IP could be promoted to this position, which of course would lead to undesirable results such as

- (22)a. *I think [yesterday, [that [Bill left the country e_i]]]
 b. *I think [the country, [that [Bill left e, yesterday]]]

Let us thus assume that once there is no overt movement to the pre-IP position as enforced by cyclic WH-movement, SpecCP will be absent and CP gets identified with C'. There is now a straightforward explanation of the **that**-trace effect which does not involve the MC. (For this purpose let us again assume the strong version of the VMH, as suggested in George (1980), i.e. in the context of VM syntactic movement must not apply). Given our assumptions about pruning of SpecCP in cases in which no WH-operator will access it, it follows from the strong VMH that in a VM-context no movement will take place and SpecCP will be pruned. Consider the case in (6), but this time without an intermediate trace in SpecCP:

- (23) Who_i do you [₁ p e_i [₁ p believe [_{CP=C'} that [_{IP} e_i [would win]]]

According to the assumptions of the Barriers framework, the CP inherits barrierhood from IP. Following the definition of Barrier given in (4) above, IP is not a barrier (see (4b)), but by (4a) a category that immediately dominates it can inherit barrierhood from it. This is in our case the CP. Notice that according to the definition in (3), IP is a BC for the trace in subject position, because it dominates the trace and it is not L-marked. Here we derive an ECP-violation: The subject-trace is neither θ -governed nor (locally) antecedent governed, and the sentence is correctly ruled out. The idea is very simply this: Once the VMH applies in a simplex CP, requiring that the subject of the contained IP does not undergo movement, this CP's specifier remains inert and will therefore be pruned. When syntactic long movement applies in such a situation, it cannot use a specifier position as an escape hatch. Therefore no intermediate trace can be left. Therefore within CP, the subject trace will not be antecedent governed.⁶ We can thus account for the **that**-trace effect in English without invoking the MC.

It should be noticed that a similar analysis of the **that**-trace effect has been proposed in the 'On Binding' framework by Taraldsen (1980). Taraldsen suggested that the trace in COMP, which then was a single position, is adjoined to the right of the complementizer. In the case of subject extraction this leads to a sequence [_{COMP} that t_i] t_i, which is barred by the VMH. Other things being equal, a subject-trace in such a context remains ungoverned. Two remarks have to be made on this: First, within the generalized X'-schema of the Barriers framework, the sequence [_{COMP} that t_i] is impossible, at least for languages like English and German. Second, there is no evidence that these languages have ever had the sequence COMP + WH. Instead earlier stages of English had the sequence WH + COMP, and German and Dutch dialects still retain such instances. It is thus more plausible to argue that due to the applicability of the VMH in English a potential antecedent governor for the subject trace does not reach SpecCP.

Let us now turn to German: We have already indicated in section 2 that in German as well as in Dutch the **that**-trace effect is absent. Claims to the contrary seem to be built on a spurious generalization about certain stylistic factors and aspects of prescriptive grammar. At any rate, speakers of German who reject the extraction of subjects from **daß**-clauses will mostly also reject the extraction of objects. A direct application of the ECP as classically stated would fail to explain this. For more discussion of these issues see Bayer (1984). Consider now the grammatical sentence in (24).

(24) Wer glaubst du [daß gewinnen würde]?

"*Who do you think that would win?"

Under the assumption that the VM does not hold in German and that WH-elements are obligatorily moved to a pre-IP position, it follows that (24) has the representation in (25):

(25) Wer, glaubst du [_{CPE} [_C daß [_{IPE} gewinnen würde]]]

The difference between German and English is constituted by the fact that German involves an intermediate trace in SpecCP while English does so only when a category is moved non-vacuously within CP.

This brings us back to adjunct movement. As one can expect from the present discussion, adjunct movement is grammatical in German. As (5) in section 2 showed, it is also grammatical in English. We have pointed out that the MC creates a problem for adjunct movement, because an overt complementizer could act as an intervener preventing the adjunct trace from being antecedent governed. In our account such a problem cannot arise. Notice that adjuncts move non-vacuously in English, as seen in cases like **How did Bill fix his car?** Consequently, the adjunct will leave an intermediate trace in SpecCP, and no ECP-problem will arise:

(26) How, do you think [_{CPE} [_C that [_{IP} Bill fixed his car e_i]]]

It will be noticed that this approach is close in spirit to Koopman (1983). Koopman suggests that the ECP-effects arise in English because of the asymmetry between "do-supported"

object-extraction and non-supported subject-extraction. Working in the classical GB-theory with a single branching COMP-node, however, it is not clear how this intuition could meet with the theoretical machinery at hand. Koopman suggests that the difference between (27a) and (27b)

- (27)a. * $[S'[_{COMP} \text{ who, did}] [e, \text{leave}]]$ (**did** unstressed)
 b. $[S'[_{COMP} \text{ who, did}] [\text{John see } e_i]]$

arises because of the following: While the WH-item is prevented from antecedent governing the trace by the intervening **did** in both cases, (27b) is rescued by means of theta- (or lexical) government. Naturally, this proposal faces the same difficulties with adjunct-movement as Chomsky's. (28) should be ungrammatical, because of an ECP-violation:

- (28) $[S'[_{COMP} \text{ how, did}] [\text{Bill fix the car } e_i]]$

Given the independently motivated phrase structure for CP as proposed in Barriers, it is even less clear how a filled C-position could block antecedent government. Another problem for Koopman's approach arises when negation comes into play. Notice that the negated version of (27a) is grammatical.

- (29) $[S'[_{COMP} \text{ who, didn't}] [e, \text{leave}]]$

Assuming that **didn't** appears in C-position, it is unlikely that (27a) should be ruled out along the lines of a **that**-trace violation.

According to the VMH there is simply no reason to apply I-to-C movement in a case like (27a). At the same time, Koopman's insight is met that Dutch does not develop ECP-effects, because there is no subject/object asymmetry with respect to the equivalent of **do**-support, namely V2.

Two other questions will still have to be considered. The first is: Will the ECP be "over-satisfied" in the case of object-movement i.e., is an object-trace in the same way theta-governed and antecedent governed? The second question is: How do we account for the grammaticality of subject-movement in English when the complementizer is absent? I shall postpone an answer to the first question until we begin to deal with WH-island violations in German. For the time being one should keep in mind that θ -government may in the end play a much weaker role than originally assumed.

Now to the second question: We have assumed that by virtue of the VMH a WH-subject does not move within CP in English. Imagine now that in a case such as (30)

- (30) Who did you believe would win?

there is an empty complementizer. Chomsky (1986) assumes exactly that. Why does the MC not rule out such cases? The answer is that an empty complementizer will not count as a minimality barrier, because it does not have features. Our own proposal would need some artificial stretching if we want to derive the grammaticality of (30) along similar lines. Notice that without any further stipulation, our account would simply rule out (30) in the same way as it rules out **that**-trace examples. Let us now assume instead that (30) does

not contain more structure than it shows, and see how far we can get with this minimalistic assumption. The complement embedded by the verb **believe** is then a bare IP, as shown in (31).

(31) Who_i did you [_{CP} e_i [_{IP} believe [_{IP} e_i would win]

From this analysis no problem can arise for antecedent government. Problems may arise, however, for the proper analysis of infinitival complements and for the level of LF. Notice first that the verb **believe** in (31) can in principle govern the subject-trace in the same way as it governs the subject-NP in cases of exceptional case marking (ECM), e.g. **She believes him to be a liar**. The WH-word in (31), however, is a clear nominative. If the complement is as unprotected as shown in (31), the case can not be clearly distinguished from an ECM-case. One possible solution could be that the trace is also case-marked by the I of its clause. Thus, it cannot get case-marked again. Another solution would be to say that IP is still protected by CP, but that in the presence of a zero-complementizer and a VM-context the category CP prunes down to IP. This gives us the equation $CP = C' = IP$. That something like this is indeed necessary is suggested by a second consideration. Notice that in a sentence such as **I didn't ask who informed Bill** the WH-word must be able to access SpecCP at the level of LF in order to fulfill the selection requirements of the matrix verb. If nothing but IP were present, WH would have to be adjoined in the fashion of quantified NPs. In languages which structurally provide for an operator position this appears unnatural. Notice also that we never find cases such as ***that whom Bill likes**, ***that whose book Bill likes**, etc. The only alternative is to still propose a CP at the level of LF. In this CP an operator position can be accessed at LF by elements which are not moved in the syntax due to the VMH. On the other hand, what counts for the ECP and antecedent government must be the structure lacking a distinct category CP or C'. Let us therefore adopt the following stipulation:

(32) **Condition on zero-categories**

Phonetically absent heads cannot induce barriers for movement or antecedent government

We will shortly see why the notion of antecedent government has to be included in this restriction on barriers. It is easy to see that antecedent government by an intermediate trace reduces to movement. Thus I will not further elaborate on this here. Example (31) will then involve the following substructure:

(33) ...believe [_{CP=C'=IP} e, would win]

If CP is not a barrier, direct access to the subject-trace is possible; on the other hand, the structure retains the potential of forming an operator/variable structure at LF.

Support for the correctness of this view is provided by Haider (1988). What Haider suggests is that – contrary to much work in GB-syntax – projections do not involve empty heads. Functional categories (such as I and C) can appear as “secondary specifications” on categories with lexical content. Haider says that “a matching projection is a projection superimposed on an existing projection such that the nodes of the primary projection serve

as secondary nodes of the superimposed projection" (p.112). Our example (31) will then retain a CP and C' in the matrix clause, but map the lower CP into the contained IP. According to this, (31) will look as follows:⁷

(31') [_{CP}Who, [_{C'}did [_{IP}you [_iPE, [_iPbelieve [_{CP\I}PE, [_{C'\I'}would [win]]]]]]]]]

According to (32), the two traces displayed are not separated by a barrier. On the other hand, we can retain the option of not having PRO governed (by the matrix verb) in cases of control. For the time being, let us assume that infinitives with a controlled PRO or PRO_{arb} count as CPs with respect to matching projections, while ECM-infinitives are never more than IPs.⁸

In sum, to the extent that the latter proposals are tenable, we have found a simple alternative to the explanation of the that-trace effect in Chomsky (1986). Chomsky tries to account for the effect with the Minimality Condition (MC). We have instead shown that it follows quite naturally when we assume that the VMH has a different outcome in English than in German or Dutch. Under the VMH it is also possible to avoid problems which arise from an MC-account of adjunct traces in English. What seems to play a key role in our account is antecedent government. In order to see more clearly which role it plays, let us now turn to WH-islands in English and Dutch/German.

4. WH-islands

In this section we will contrast WH-islands in English and German. It is well known, at least since Huang (1982), that WH-island violations in English are tolerable when a lexically governed element moves out, but not when subjects or adjuncts move. The crucial data are given in (34):

- (34)a. ?What, do you wonder [how, John fixed e, e,]?
 b. *How, do you wonder [what, John fixed e, e,]?
 c. *Who, do you wonder [what, e, fixed e,]?

Huang's explanation is that object extraction is licit because lexical government fulfills one half of the ECP, while cases like (34b) and (34c) are ECP-violations, since neither antecedent government nor lexical government holds.⁹ Consider now similar WH-island violations in German:

- (35)a. *Was, hat er sich gefragt [_{CP} wie, [Fritz e, e, repariert hat]]?
 what did he (refl) ask how Fritz repaired has
 "For which thing x did he ask himself how Fritz had repaired x?"
 b. *Wie, hat er sich gefragt [_{CP} was, [Fritz e, e, repariert hat]]?
 "For which manner y did he ask himself what Fritz had repaired in y?"
 c. *Wer, hat er sich gefragt [_{CP} was, [e, e, repariert hat]]?
 who did he (refl) ask what repaired has

“For which person x did he ask himself what x had repaired”

In contrast to the English examples in (34), the German examples are all ungrammatical. According to van de Koot (1988) the same holds for Dutch. As in English we can assume that antecedent government is not available, but that lexical government is somehow too weak in German to rescue cases of object extraction. This seems appropriate as a first generalization, but it is certainly too vague to have any theoretical significance. In order to see that the situation in German is even more complicated than suggested by the pattern in (35), consider the examples in (36) below. For German, examples of this kind were detected and first discussed in Fanselow (1987).

- (36)a. ?Radios, weiß ich nicht [wer e, repariert]
radios know I not who repairs
“As for radios, I don’t know who repairs (them)”
- b. *Linguisten, weiß ich nicht [was, e, reparieren]
linguists know I not what repair
“As for linguists, I don’t know what (they) repair”

As (36a) shows, it is not in principle impossible to extract from a WH-island in German when the material to be extracted is [-WH]. There seem to be other restrictions as well. For example, both examples in (36) involve bare plurals, i.e., typical topic phrases. When they are replaced by definite NPs like *dieses Radio* (“this radio”) relatively well-formed examples become less acceptable. At least, it needs some stretching to make the examples acceptable. What are topic phrases? A rough and ready characterization seems to be that they are discourse elements which are syntactically unconnected to the clause which they introduce. So there has to be an element in the clause which serves as a link between the topic phrase and a position in the clause to which it will correspond semantically. According to certain analyses such a link may be a pronominal or an empty operator. Let us pursue this idea, and suggest that we are dealing here with topic constructions in which it is not the first element which undergoes movement, but rather an empty operator in the sense of Chomsky (1986), where such an operator is motivated for the treatment of parasitic gap constructions. If this is correct, the structure of (36a) would rather be as in (37):

- (37) ?Radios, O, weiß ich nicht [wer e, repariert]

In (37) O is an empty operator which is coindexed with the topic phrase *Radios*.¹⁰ Empirical support for this analysis is gained by the fact that the results are less acceptable when elements appear in topic position which cannot normally be pronominalized. Consider the following contrast:

- (38)a. ?Mit Saxophonen weiß ich nicht wer Bach gut spielen kann
with saxophones know I not who Bach well play can
- b. ??Ohne Oboen weiß ich nicht wer die Matthäus-Passion gut spielen kann
without oboes know I not who the M-P well play can

- c. ?*In dieser Art weiß ich nicht wer Bach spielt
 in this way know I not who Bach plays

The *mit*-PP in (38a) can be pronominalized with *damit* (“therewith”), while there is no specific pronominal adverb for *ohne*-PPs in German. It is, however, marginally possible to refer to it with the neutral *da*. For manner-PPs such as *in dieser Art* even this is impossible. This situation is mirrored in the decreasing grammaticality of the examples of (38). Notice that adjuncts appear now in a different perspective. One could argue that all the PPs involved in (38) have adjunct status. How is it then possible that some are more acceptable than the others? If it were primarily lexical government that matters, the difference would be difficult to explain. Furthermore it would be unclear why the extraction of [+WH] elements from WH-islands in German is in general impossible. We would rather expect a clear asymmetry between subjects and adjuncts on one side and L-marked objects on the other, as in English.

We have assumed above that WH-islands are the result of the lower WH-element being in SpecCP, thus preventing a trace in the complement from being antecedent governed. This appears to be the right conclusion for English, but we will argue below that it might not be desirable for German. Ultimately, our arguments will carry over to English as well. In order to show this, we will first have a closer look at “COMP” in German.

4.1 WH as a complementizer

The Barriers framework draws a sharp distinction between complementizers and WH-elements which move to the SpecCP-position. Chomsky (1986:4 and 71) following Travis (1984) and earlier work by Emonds proposes the ‘Head Movement Constraint’ (HMC), which essentially says that only X^0 -categories can move to a head position Y^0 that governs XP.¹¹ Since C is a head position and since WH-phrases are generally X^{max} , it follows that WH-phrases cannot move to C. When we consider South-German examples such as (17) of section 3, which we repeat here for convenience, the HMC seems to be the appropriate generalization:

- (17)a. Arabella weiß [wer [daß [den Hans gesehen hat]]]
 b. die Frau [die [wo [den Hans gesehen hat]]]

In Bayer (1984) I have shown, however, that this is but one option for complementation. The alternative option was that the complementizer may be missing in D-structure and that the WH-element will move precisely in the place otherwise occupied by the complementizer. My main argument in favor of this analysis was that in Bavarian the C-position obligatorily inflects for second person and number, i.e. irrespective of whether it contains the finite verb or the complementizer. Observe the following pattern, (which also holds for relative clauses):

- (39)a. Ea hot g’sagt [du [wui-st [net mi:tmacha]]]
 he has said you want-(2sg) not with-do
 “He said you would not like to join”

- b. Ea hot g'sagt [e:s [woi-ts [net mi:tmacha]]]
(2pl)
- c. Ea hot g'sagt [dass-st [du net mi:tmacha wui -st]]
he has said that-(2sg) you not with-do want-(2sg)
- d. Ea hot g'sagt [dass-ts [e:s net mi:tmacha woi-ts]]
(2pl) (2pl)

As was shown in Bayer (1984) there are good reasons not to assign normal clitic status to the morphemes *-st* and *-ts*, but rather the status of person/number inflection as it occurs on the verb. When WH-phrases move across a complementizer to SpecCP ("COMP2" of Bayer, 1984), it is still the complementizer which carries the inflection, not the WH-phrase: **wann-st* *dass-st* or **wann-st* *dass* are ungrammatical. Observe now the data in (40), where the complementizer is missing and where it is the WH-phrase itself which carries the inflection:

- (40)a. Ea hot g'frogt [[wos fia Schuach]-st [du o:ziag-st]]
he has asked what for shoes-(2sg) you put on
"He asked what kind of shoes you will put on"
- b. Ea hot g'frogt [[wos fia Schuach]-ts [e:s o:ziag-ts]]
(2pl) (2pl)

This can be taken as evidence that it is indeed the C-position which hosts the WH-phrase at S-structure. Of course, there is always the possibility that what we see is not really cliticization of the 2nd person morpheme to C, but to SpecCP, across an empty C-position. Given the fact, however, that this morphology is precisely the same that is also carried by I (which does move to C according to the HMC), it is not very plausible that things should be different in cases such as (40). Thus far, this is not really a conclusive argument, but there are others which support this conclusion. Grewendorf (1988:ch.11) gives a review of which I want to mention two arguments both of which appear in Reis (1985). The first has to do with the verb-second (V2) phenomenon. Reis points to a number of serious problems for an analysis of German clause structure which suggests a single underlying CP-structure (in her terminology S''). One such problem is how to prevent the finite verb from moving to the C-position in embedded clauses. The widely shared generalization is that V2 applies whenever the C-position is empty. For embedded clauses, however, V2 is sharply ungrammatical. One of Reis' examples is given in (41):

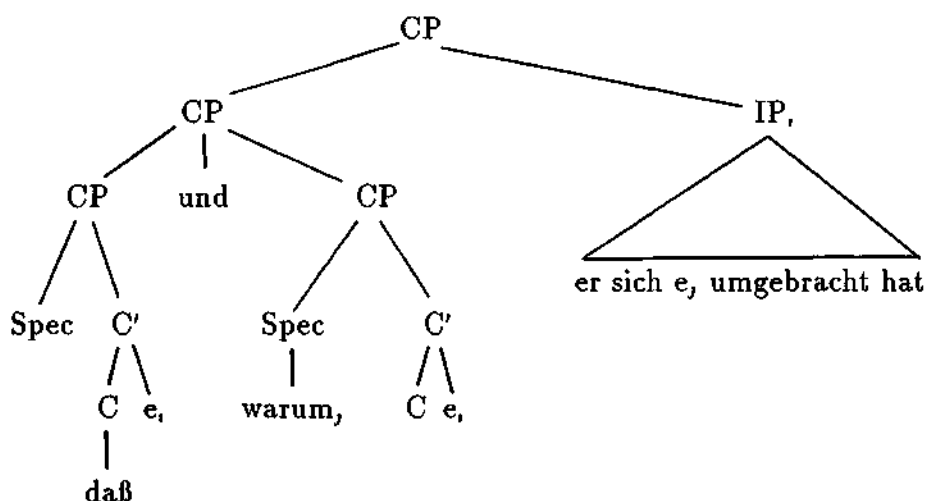
- (41)a. Weiß der Himmel wie man diese Fakten erklären kann
knows the heaven how one these facts explain can
"God knows how to explain these facts!"
- b. *Weiß der Himmel wie kann man diese Fakten erklären

An analysis which moves WH to C can avoid the overgeneration in (41b) without further stipulation. Reis also observes that coordination creates a problem. Notice the following examples:

- (42)a. Ich habe erfahren daß und warum er sich umgebracht hat
 I have heard that and why he himself killed has
 "I have heard that and why he has killed himself"
- b. Ich habe erfahren daß und in wen er sich verliebt hat
 I have heard that and in who he himself be-loved has
 "I have heard that and with who he has fallen in love"

The most straightforward analysis of these cases is to coordinate the two COMPs. But this is only possible when they are in parallel positions (see Williams, 1978). Either they are both in SpecCP or both in C. The former option is clearly out, but the second one seems to be in conflict with the HMC. Another possibility, right-node raising (actually the only option left under Williams' proposal), is also problematic. As the following partial structure of (42a) shows, the first conjunct will crucially have to involve a SpecCP-position:

(43)



The SpecCP-position in the first conjunct, however, is both unmotivated and prone to invite other problems, as we have already argued in connection with (22) in section 3.

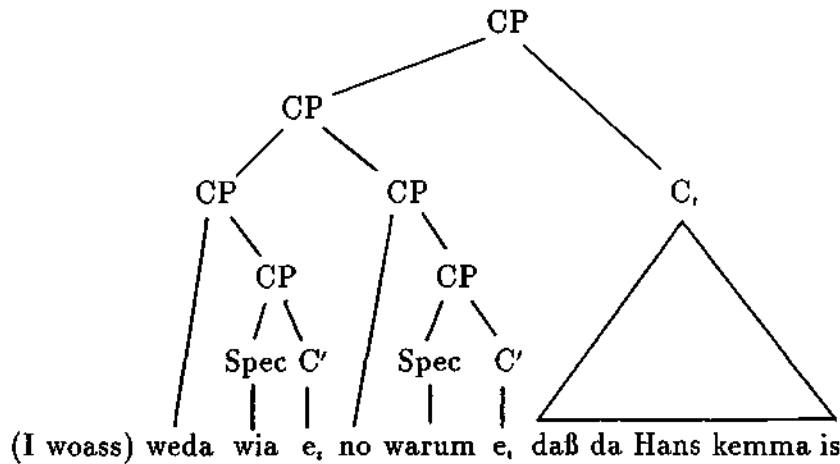
In Bavarian, obligatory movement of WH to SpecCP makes the wrong predictions for coordination even more clearly. So-called "double COMP" clauses are predicted to coordinate with "single COMP" clauses. But as the examples in (44c,d) show, this prediction is not borne out:

- (44)a. I woass weda wia no warum da Hans kemma is
 I know neither how nor why Hans come has
 "I know neither how nor why Hans has come"

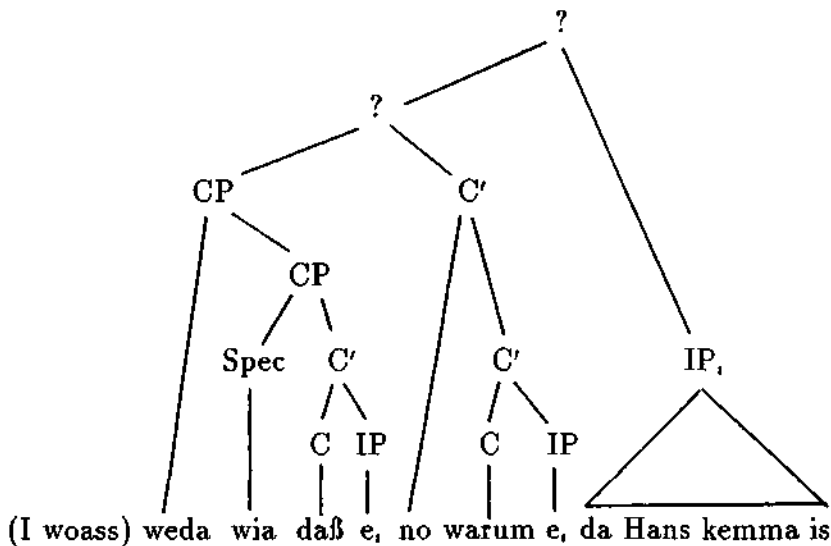
- b. I woass weda wia no warum daß da Hans kemma is
that
- c. *I woass weda wia daß no warum da Hans kemma is
- d. *I woass weda daß no warum daß da Hans kemma is

The examples (44a,b) show strict parallelism in the material to be coordinated. In (44a) only one WH-element appears per position. In (44b) the coordination is again parallel, leaving the **daß**-clause to the right of the WH-elements. (44c,d), however, deviate from this parallelism. The following representations of (44b) and (44c) with right-node raising show in which way:

(44b')

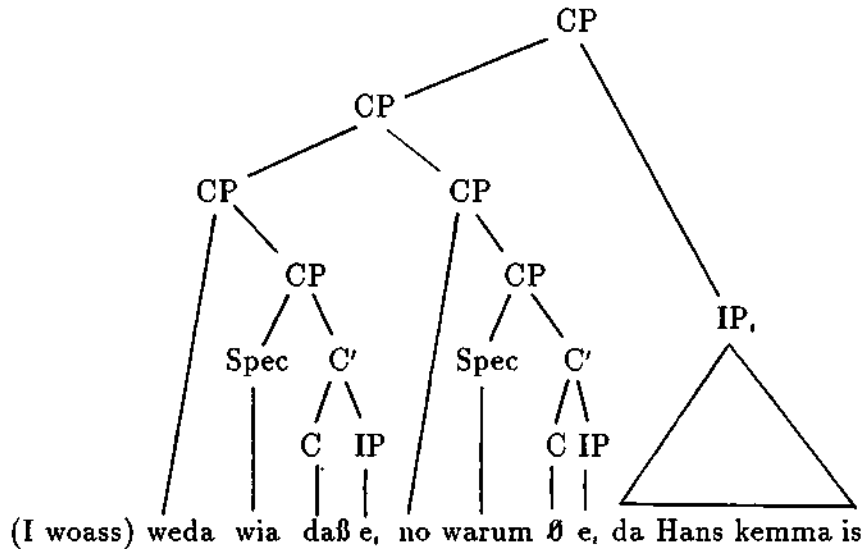


(44c')



Moving **warum** to SpecCP instead of moving it to C would lead to the following structure of (44c):

(44c'')



There is presumably nothing in (44c'') which would invoke a known principle to explain its ungrammaticality. Under the assumption, however, that WH-elements can move to C, the non-parallelism in the conjunct follows immediately, as depicted by the representation in (44c').

Another problem for WH-to-SpecCP in the absence of an overt complementizer emerges from the following examples of a WH-imperative, a construction type studied by Reis and Rosengren (1988). The following example is taken from their work.

- (45)a. Wem, sag mal [e, daß du die Rezension e, anvertraut hast]
 who say ? that you the review given have
 "Who have you given the job of doing the review, tell me!"
 b. *Wem, sag mal [e, du die Rezension e, anvertraut hast]

Under the standard approach, the trace would in both cases be in SpecCP. As a consequence, the ungrammaticality of (45b) would have to be accounted for by some novel principle. When the WH-item in (45b), however, moves from a C-position, it is obvious that standard assumptions about WH-movement are violated: A WH-chain involving an intermediate head position would be ruled out by the requirement that it is only A'/operator-positions and one potential A-position at the foot of the chain which can be linked. According to the HMC, no syntactic phrase can ever pass through a head-position such as C. Precisely this could have happened in (45b), i.e. the example is most likely not ruled out because there is no overt complementizer present, but because WH has moved through the only syntactically available position, namely C. Note, however, that it would be paradoxical to reject the HMC on one side, namely for the purpose of moving WH to C, and invoke it again in order to prevent WH from passing through C. We shall shortly show that given one reasonable assumption, the conflict is only apparent. For the time being it is enough to witness that in all of the well-formed examples of (44) and (45) WH-movement terminates in a C-position, while in

(45b) it would just use C as an escape position. As we have already indicated in connection with the absence of V2 effects in WH-complements (see (41b)), WH in root-clauses clearly terminates in SpecCP, while it must terminate in C in embedded clauses. Otherwise V2 would apply, deriving an ungrammatical structure.

A final argument in favor of German WH-to-C and against WH-to-SpecCP in the absence of a lexical complementizer is this: Contrary to English, German systematically lacks WH-phrases initiating infinitival zu-complements; (see Tappe, 1984 who argues for a general lack of C(OMP) in German infinitives). Let us assume, following earlier work by den Besten (1977/1983), that lexical complementizers such as Dutch *dat*, *of* and German *daß*, *ob* realize features of I(NFL). Note that I-to-C movement is crucially dependent on such feature sharing. If we assume that German WH-elements never touch the C-position and C remains empty, as shown in (46b), some other principle must be invoked which rules out sentences like (46a).

- (46)a. *Die Leute wissen nicht was zu tun
 the people know not what to do
 b. Die Leute wissen nicht [_{CP} was_i [_{C'} [_{IP} e_i zu tun]]]
 c. Die Leute wissen nicht [_{CP} [_{C'} was_i [_{IP} e_i zu tun]]]

If however, WH-elements – may be as a language specific option – must realize I-features, they must be banned from the C-position of zu-infinitives in the same way as *daß* and *ob* are. The ungrammaticality of (46a) follows straightforwardly in a representation like (46c). As a cautious note, let me say that neither *that* nor WH in English seem to realize I-features. The English C-position seems to be rather neutral in this respect. This, of course, leaves the possibility that WH moves to C in English as well. We will return to this issue at the end of section 4.2.

Under the assumption that both complementizers and WH-phrases can occupy the C-position at S-structure, a whole set of problems vanishes immediately, – problems which may find an account otherwise, but for which it is highly unlikely that a simpler and more elegant account can be found than the one sketched above: The absence of V2 in embedded WH-clauses, CP-coordination with subsequent right-node raising, the absence of WH-chains without an unfilled C-slot, as well as the lack of WH in zu-infinitives are naturally explained by the assumption that WH-elements in German bear I-features and that they are moved to C in syntax. Let us therefore seriously adopt this approach and ask what its further consequences are.

Two problems can be witnessed: The first has to do with the status of the HMC, the second with the representation of WH-operators in LF. What I have to suggest in order to retain the advantages of the HMC is that WH-phrases may be made visible by virtue of their feature [+WH] and not by their being phrasal or not.¹² Under this natural assumption WH-phrases can realize the head of CP in the absence of a lexical complementizer. Let us state this in the form of a condition on C-realization:

(47) Condition on C-realization

- (i) If the C-position is not filled by a lexical complementizer at D-structure, C will be derivationally realized by a verb marked [+I] or by a phrase marked [+WH].
- (ii) [+WH]-marked elements count by their WH-feature as head categories for the HMC.

Provided (47) can cope with the requirements of the HMC, let us now see how WH-operators can be represented at LF. Following Chomsky and many others, we will assume that WH-elements must have access to an operator position for reasons of semantic interpretation. In languages without overt WH-movement it is assumed that an operator position is only accessed in LF. Does that mean that these languages have a fully operative C-projection system in their syntax which happens to never get used? This seems rather awkward. I would like to suggest instead that in these languages a structural operator position is only present at LF. Recall that the same problem was already discussed at the end of section 3. There it was suggested that subject-extraction from headless finite complements was possible in English because CP is syntactically identified with IP. This prevents syntactic movement from accessing landing sites within CP, but it does not prevent LF-movement. That it cannot do so is indirectly shown by Baker's (1970) study of multiple questions in English. In a case like Baker's example **Who remembers where we bought which book?** the lowest WH-phrase can be paired up with the highest WH-phrase giving the reading "For which person x and for which book y, x remembers at which place z we bought y". Clearly, English does not permit this to take place in the syntax in the same way as it does not permit cyclic WH-movement out of WH-islands. I take this as evidence that LF-movement operates independently of a syntactic SpecCP-position. For our analysis of WH in C-position this means that CP will be identified with C', thus closing the complement for further syntactic movement, while retaining the property of getting the WH-in-C into an operator position. The following condition seems appropriate for both German and English:

(48) Condition on WH-interpretation

A WH-phrase which realizes C by (47) will realize an operator position at LF.¹³

In the next section we will have a closer look at what this system can do for the explanation of WH-island effects in English and German.

4.2 Extraction from WH-islands?

Recall first the English and German data presented in (34) and (35) respectively, which we repeat below:

- (34)a. ?What_i do you wonder [how_j John fixed e_i e_j]?
 - b. *How_j do you wonder [what_i John fixed e_i e_j]?
 - c. *Who_i do you wonder [what_j e_i fixed e_j]?
 - d. *How_j do you wonder [who_i e_j fixed e_i]?
 - e. *Who_i do you wonder [how_j e_j fixed e_i]?
 - f. *How_j do you wonder [who_i e_j fixed e_i]?
 - g. *Who_i do you wonder [how_j e_j fixed e_i]?
 - h. *How_j do you wonder [who_i e_j fixed e_i]?
 - i. *Who_i do you wonder [how_j e_j fixed e_i]?
 - j. *How_j do you wonder [who_i e_j fixed e_i]?
 - k. *Who_i do you wonder [how_j e_j fixed e_i]?
 - l. *How_j do you wonder [who_i e_j fixed e_i]?
 - m. *Who_i do you wonder [how_j e_j fixed e_i]?
 - n. *How_j do you wonder [who_i e_j fixed e_i]?
 - o. *Who_i do you wonder [how_j e_j fixed e_i]?
 - p. *How_j do you wonder [who_i e_j fixed e_i]?
 - q. *Who_i do you wonder [how_j e_j fixed e_i]?
 - r. *How_j do you wonder [who_i e_j fixed e_i]?
 - s. *Who_i do you wonder [how_j e_j fixed e_i]?
 - t. *How_j do you wonder [who_i e_j fixed e_i]?
 - u. *Who_i do you wonder [how_j e_j fixed e_i]?
 - v. *How_j do you wonder [who_i e_j fixed e_i]?
 - w. *Who_i do you wonder [how_j e_j fixed e_i]?
 - x. *How_j do you wonder [who_i e_j fixed e_i]?
 - y. *Who_i do you wonder [how_j e_j fixed e_i]?
 - z. *How_j do you wonder [who_i e_j fixed e_i]?
 - aa. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ab. *How_j do you wonder [who_i e_j fixed e_i]?
 - ac. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ad. *How_j do you wonder [who_i e_j fixed e_i]?
 - ae. *Who_i do you wonder [how_j e_j fixed e_i]?
 - af. *How_j do you wonder [who_i e_j fixed e_i]?
 - ag. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ah. *How_j do you wonder [who_i e_j fixed e_i]?
 - ai. *Who_i do you wonder [how_j e_j fixed e_i]?
 - aj. *How_j do you wonder [who_i e_j fixed e_i]?
 - ak. *Who_i do you wonder [how_j e_j fixed e_i]?
 - al. *How_j do you wonder [who_i e_j fixed e_i]?
 - am. *Who_i do you wonder [how_j e_j fixed e_i]?
 - an. *How_j do you wonder [who_i e_j fixed e_i]?
 - ao. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ap. *How_j do you wonder [who_i e_j fixed e_i]?
 - aq. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ar. *How_j do you wonder [who_i e_j fixed e_i]?
 - as. *Who_i do you wonder [how_j e_j fixed e_i]?
 - at. *How_j do you wonder [who_i e_j fixed e_i]?
 - au. *Who_i do you wonder [how_j e_j fixed e_i]?
 - av. *How_j do you wonder [who_i e_j fixed e_i]?
 - aw. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ax. *How_j do you wonder [who_i e_j fixed e_i]?
 - ay. *Who_i do you wonder [how_j e_j fixed e_i]?
 - az. *How_j do you wonder [who_i e_j fixed e_i]?
 - ba. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bb. *How_j do you wonder [who_i e_j fixed e_i]?
 - bc. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bd. *How_j do you wonder [who_i e_j fixed e_i]?
 - be. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bf. *How_j do you wonder [who_i e_j fixed e_i]?
 - bg. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bh. *How_j do you wonder [who_i e_j fixed e_i]?
 - bi. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bj. *How_j do you wonder [who_i e_j fixed e_i]?
 - bk. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bl. *How_j do you wonder [who_i e_j fixed e_i]?
 - bm. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bn. *How_j do you wonder [who_i e_j fixed e_i]?
 - bo. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bp. *How_j do you wonder [who_i e_j fixed e_i]?
 - bq. *Who_i do you wonder [how_j e_j fixed e_i]?
 - br. *How_j do you wonder [who_i e_j fixed e_i]?
 - bs. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bt. *How_j do you wonder [who_i e_j fixed e_i]?
 - bu. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bv. *How_j do you wonder [who_i e_j fixed e_i]?
 - bw. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bx. *How_j do you wonder [who_i e_j fixed e_i]?
 - by. *Who_i do you wonder [how_j e_j fixed e_i]?
 - bz. *How_j do you wonder [who_i e_j fixed e_i]?
 - ca. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cb. *How_j do you wonder [who_i e_j fixed e_i]?
 - cc. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cd. *How_j do you wonder [who_i e_j fixed e_i]?
 - ce. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cf. *How_j do you wonder [who_i e_j fixed e_i]?
 - cg. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ch. *How_j do you wonder [who_i e_j fixed e_i]?
 - ci. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cj. *How_j do you wonder [who_i e_j fixed e_i]?
 - ck. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cl. *How_j do you wonder [who_i e_j fixed e_i]?
 - cm. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cn. *How_j do you wonder [who_i e_j fixed e_i]?
 - co. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cp. *How_j do you wonder [who_i e_j fixed e_i]?
 - cq. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cr. *How_j do you wonder [who_i e_j fixed e_i]?
 - cs. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ct. *How_j do you wonder [who_i e_j fixed e_i]?
 - cu. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cv. *How_j do you wonder [who_i e_j fixed e_i]?
 - cw. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cx. *How_j do you wonder [who_i e_j fixed e_i]?
 - cy. *Who_i do you wonder [how_j e_j fixed e_i]?
 - cz. *How_j do you wonder [who_i e_j fixed e_i]?
 - da. *Who_i do you wonder [how_j e_j fixed e_i]?
 - db. *How_j do you wonder [who_i e_j fixed e_i]?
 - dc. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dd. *How_j do you wonder [who_i e_j fixed e_i]?
 - de. *Who_i do you wonder [how_j e_j fixed e_i]?
 - df. *How_j do you wonder [who_i e_j fixed e_i]?
 - dg. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dh. *How_j do you wonder [who_i e_j fixed e_i]?
 - di. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dj. *How_j do you wonder [who_i e_j fixed e_i]?
 - dk. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dl. *How_j do you wonder [who_i e_j fixed e_i]?
 - dm. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dn. *How_j do you wonder [who_i e_j fixed e_i]?
 - do. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dp. *How_j do you wonder [who_i e_j fixed e_i]?
 - dq. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dr. *How_j do you wonder [who_i e_j fixed e_i]?
 - ds. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dt. *How_j do you wonder [who_i e_j fixed e_i]?
 - du. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dv. *How_j do you wonder [who_i e_j fixed e_i]?
 - dw. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dx. *How_j do you wonder [who_i e_j fixed e_i]?
 - dy. *Who_i do you wonder [how_j e_j fixed e_i]?
 - dz. *How_j do you wonder [who_i e_j fixed e_i]?
 - ea. *Who_i do you wonder [how_j e_j fixed e_i]?
 - eb. *How_j do you wonder [who_i e_j fixed e_i]?
 - ec. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ed. *How_j do you wonder [who_i e_j fixed e_i]?
 - ee. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ef. *How_j do you wonder [who_i e_j fixed e_i]?
 - eg. *Who_i do you wonder [how_j e_j fixed e_i]?
 - eh. *How_j do you wonder [who_i e_j fixed e_i]?
 - ei. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ej. *How_j do you wonder [who_i e_j fixed e_i]?
 - ek. *Who_i do you wonder [how_j e_j fixed e_i]?
 - el. *How_j do you wonder [who_i e_j fixed e_i]?
 - em. *Who_i do you wonder [how_j e_j fixed e_i]?
 - en. *How_j do you wonder [who_i e_j fixed e_i]?
 - eo. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ep. *How_j do you wonder [who_i e_j fixed e_i]?
 - eq. *Who_i do you wonder [how_j e_j fixed e_i]?
 - er. *How_j do you wonder [who_i e_j fixed e_i]?
 - es. *Who_i do you wonder [how_j e_j fixed e_i]?
 - et. *How_j do you wonder [who_i e_j fixed e_i]?
 - eu. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ev. *How_j do you wonder [who_i e_j fixed e_i]?
 - ew. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ex. *How_j do you wonder [who_i e_j fixed e_i]?
 - ey. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ez. *How_j do you wonder [who_i e_j fixed e_i]?
 - fa. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fb. *How_j do you wonder [who_i e_j fixed e_i]?
 - fc. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fd. *How_j do you wonder [who_i e_j fixed e_i]?
 - fe. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ff. *How_j do you wonder [who_i e_j fixed e_i]?
 - fg. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fh. *How_j do you wonder [who_i e_j fixed e_i]?
 - fi. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fj. *How_j do you wonder [who_i e_j fixed e_i]?
 - fk. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fl. *How_j do you wonder [who_i e_j fixed e_i]?
 - fm. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fn. *How_j do you wonder [who_i e_j fixed e_i]?
 - fo. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fp. *How_j do you wonder [who_i e_j fixed e_i]?
 - fq. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fr. *How_j do you wonder [who_i e_j fixed e_i]?
 - fs. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ft. *How_j do you wonder [who_i e_j fixed e_i]?
 - fu. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fv. *How_j do you wonder [who_i e_j fixed e_i]?
 - fw. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fx. *How_j do you wonder [who_i e_j fixed e_i]?
 - fy. *Who_i do you wonder [how_j e_j fixed e_i]?
 - fz. *How_j do you wonder [who_i e_j fixed e_i]?
 - ga. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gb. *How_j do you wonder [who_i e_j fixed e_i]?
 - gc. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gd. *How_j do you wonder [who_i e_j fixed e_i]?
 - ge. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gf. *How_j do you wonder [who_i e_j fixed e_i]?
 - gg. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gh. *How_j do you wonder [who_i e_j fixed e_i]?
 - gi. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gj. *How_j do you wonder [who_i e_j fixed e_i]?
 - gk. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gl. *How_j do you wonder [who_i e_j fixed e_i]?
 - gm. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gn. *How_j do you wonder [who_i e_j fixed e_i]?
 - go. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gp. *How_j do you wonder [who_i e_j fixed e_i]?
 - gq. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gr. *How_j do you wonder [who_i e_j fixed e_i]?
 - gs. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gt. *How_j do you wonder [who_i e_j fixed e_i]?
 - gu. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gv. *How_j do you wonder [who_i e_j fixed e_i]?
 - gw. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gx. *How_j do you wonder [who_i e_j fixed e_i]?
 - gy. *Who_i do you wonder [how_j e_j fixed e_i]?
 - gz. *How_j do you wonder [who_i e_j fixed e_i]?
 - ha. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hb. *How_j do you wonder [who_i e_j fixed e_i]?
 - hc. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hd. *How_j do you wonder [who_i e_j fixed e_i]?
 - he. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hf. *How_j do you wonder [who_i e_j fixed e_i]?
 - hg. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hh. *How_j do you wonder [who_i e_j fixed e_i]?
 - hi. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hj. *How_j do you wonder [who_i e_j fixed e_i]?
 - hk. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hl. *How_j do you wonder [who_i e_j fixed e_i]?
 - hm. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hn. *How_j do you wonder [who_i e_j fixed e_i]?
 - ho. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hp. *How_j do you wonder [who_i e_j fixed e_i]?
 - hq. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hr. *How_j do you wonder [who_i e_j fixed e_i]?
 - hs. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ht. *How_j do you wonder [who_i e_j fixed e_i]?
 - hu. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hv. *How_j do you wonder [who_i e_j fixed e_i]?
 - hw. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hx. *How_j do you wonder [who_i e_j fixed e_i]?
 - hy. *Who_i do you wonder [how_j e_j fixed e_i]?
 - hz. *How_j do you wonder [who_i e_j fixed e_i]?
 - ia. *Who_i do you wonder [how_j e_j fixed e_i]?
 - ib. *How_j do you

- (35)a. *Was, hat er sich gefragt [_{CP} wie, [Fritz e, e, repariert hat]]?
 what did he (refl) ask how Fritz repaired has
 “For which thing x did he ask himself how Fritz had repaired x?”
- b. *Wie, hat er sich gefragt [_{CP} was, [Fritz e, e, repariert hat]]?
 “For which manner y did he ask himself what Fritz had repaired in y?”
- c. *Wer, hat er sich gefragt [_{CP} was, [e_i e, repariert hat]]?
 who did he (refl) ask what repaired has
 “For which person x did he ask himself what x had repaired”

These data have pointed to the following picture: German (as well as Dutch) never allows WH-phrases to move out of WH-islands, while English marginally allows lexically governed or at least θ -governed WH-phrases to do so. Assuming that lexical government is for some reason weaker in German/Dutch than in English, this difference follows from the theory developed so far: Antecedent government will never hold, because the WH-element initiating the embedded CP is in C-position such that CP prunes down to C'. Recall that it will retain its CP-hood only for the formation of a proper operator/variable chain at LF. The complement is then closed for further syntactic movement. It can as a consequence not host an antecedent governor for the trace in IP. The only way of rescuing a trace in IP from offending the ECP is by lexical government which seems somehow insufficient in German and Dutch.

This can, however, not be the entire picture, as indicated by the German examples in (36) and (38), which we repeat here for convenience.

- (36)a. ?Radios, weiß ich nicht [wer e, repariert]
 radios know I not who repairs
 “As for radios, I don’t know who repairs (them)”
- b. *Linguisten, weiß ich nicht [was, e, e, reparieren]
 linguists know I not what repair
 “As for linguists, I don’t know what (they) repair”
- (38)a. ?Mit Saxophonen weiß ich nicht wer Bach gut spielen kann
 with saxophones know I not who Bach well play can
- b. ??Ohne Oboen weiß ich nicht wer die Matthäus-Passion gut spielen kann
 without oboes know I not who the M-P well play can
- c. ?*In dieser Art weiß ich nicht wer Bach spielt
 in this way know I not who Bach plays

Curiously, these examples cross-cut our generalization in a surprising way: With respect to WH-movement we have found that (A) WH-islands are firmly closed and (B) lexical government is not sufficient. With respect to the movement of non-WH material, i.e. according to our previous considerations movement of an empty operator we find now that (A) WH-complements are not so firmly closed and that (B) lexical government may still play a

role for the ECP in German after all. How can we account for this vexed situation? My suggestion is the following: Assume that WH-elements which realize a complementizer are fused into CP only under a special condition, namely when cyclic WH-movement applies. It is common knowledge that cyclic movement cannot pass through the C-position.¹¹ Whenever cyclic movement takes place, SpecCP must be involved. Given now the situation of a one-step moved WH-item in C, the WH-item which undergoes “long” movement has to pass through the adjacent SpecCP. This is shown in (49):

(49) $WH_1 \dots [_{CP} e_1 [_{C'} WH_2 \{_{IP} \dots e_2 \dots e_1 \dots \}]]$

I want to argue that such a situation cannot arise, because WH_1 will affect WH_2 in such a way that the latter has to move to SpecCP immediately instead of doing so only at LF. As a reason for this we assume that the feature [+WH] is shared by both WH_1 and WH_2 . Relative D-pronouns of German seem to share the WH-feature, too. Extraction of D-pronouns from WH-CPs yields firm WH-island violations which appear to be ECP violations, e.g.:

- (50)a. *Dies ist ein Patient den, ich nicht weiß wann man e, operiert hat
 This is a patient who I not know when one operated has
 “This is a patient of who I don’t know when surgery was performed on him”
 b. *Dies ist ein Patient dem, ich nicht weiß wer e, helfen könnte
 “This is a patient of who I don’t know who could help him”
 c. *Dies ist ein Patient der, ich nicht weiß warum e, operiert werden sollte
 “This is a patient of who I don’t know why he should undergo surgery”

Lexical government would predict (50c) to be ill-formed, because it is the subject which moves, but it would wrongly predict (50a,b) to be well-formed, because here it is the objects which move. Given that lexical government is too weak in German, our proposal predicts that all of these examples are ruled out by cyclic movement of the D-pronoun whatever its status is otherwise. In moving through SpecCP the D-word “pushes” the WH-item in C into the specifier position, but this makes it impossible that both WH-elements leave a trace in SpecCP. As we have argued, such an intermediate trace seems necessary, however, to satisfy the ECP in German. Let us adopt the following convention:

(51) WH-feature-sharing convention

When WH moves through SpecCP in the syntax, any WH’ which is not in situ is forwarded to SpecCP too.

(51) guarantees that WH-in-C is affected by any WH-element that moves to SpecCP. Let us now see what happens when [-WH]-items bind into WH-islands. For convenience we repeat example (36a):

- (36a) ?Radios, O, weiß ich nicht [wer e, repariert]
 “As for radios, I don’t know who repairs (them)”

Here it is an empty operator that undergoes cyclic movement. This operator presumably does not share the WH-feature with the WH-word initiating the complement. By (51), the

WH-word can remain in C-position at S-structure, and the empty operator can leave a trace in SpecCP. Thus, even when lexical government is not sufficient, this case can be accounted for by antecedent government. (36a) will be assigned the representation in (52):

(52) ?Radios, O, weiß ich nicht [_{CP}e, [_{C'}wer [_{IP}e, repariert]

Notice that this step makes the right predictions for a number of other cases which cannot be explained by lexical government or θ -government. Take adverbs and predicative adjectives, for instance.

- (53)a. Gestern früh wüßte ich nicht wer sein Auto gewaschen hat
 yesterday morning know(subj) I not who his car washed has
 "As for yesterday morning, I wouldn't know who has (then) washed his car"
 b. Dort drüben wüßte ich nicht wer sein Auto gewaschen hat over there
- (54)a. ?*So schlampig wüßte ich nicht wer sein Auto gewaschen hat
 so messily
 b. ?*Betrunken wüßte ich nicht wer sein Auto gewaschen hat
 drunk
 c. ?*Gut wüßte ich nicht wem es in deiner Gesellschaft geht
 well know(subj) I not whom(dat) it in your company goes
 "I would not know who could do well in your company"

Adverbs of time and place are clearly optional and as such unlikely to be θ -marked by the verb, (unless one assumes a strictly semantic approach in the sense of Davidson, where they would be coordinates of an associated event argument). These adverbs can easily extract from WH-islands. Manner adverbs and predicative adjectives, however, refuse to extract, as the examples in (54) show. This is independent of semantic selection. **Gut** in (54c) is strictly subcategorized by the verb **gehen**, but still extraction from the WH-clause is impossible. In order to see that this is not a **general** prohibition against extraction, consider the parallel cases in (55). Here, the prohibition against cyclic movement is absent:

- (55) a. So schlampig [glaube ich nicht [daß Hans sein Auto gewaschen hat]]
 b. Betrunken [glaube ich nicht [daß Hans sein Auto gewaschen hat]]
 c. Gut [glaube ich nicht [daß es dem Hans in deiner Gesellschaft geht]]

The same holds for separable verb-particles such as **zu-** in **zumachen** ("to close"), which in principle can undergo long movement. The list of examples could be extended. It would be difficult to argue that antecedent government from SpecCP holds in (53) and (55), but not in (54). Let us instead assume that antecedent government holds throughout, and that the difference derives from something else. We have already indicated that WH-island violations in German are likely to involve an empty operator, as it appears in (36) and (52) above. If we assume such an operator also for the examples in (53) and (54), things fall into place immediately. Notice that time and place adverbs can be pronominalized, but manner adverbs etc. cannot. This is shown in (56) and (57) below:

- (56) a. Gestern früh, **da** hat er sein Auto gewaschen
 b. Dort drüben, **da** hat er sein Auto gewaschen
- (57) a. *So schlampig, **das** hat Hans sein Auto nicht gewaschen
 b. *Betrunken, **das** hat Hans sein Auto nicht gewaschen
 c. *Gut, **das** geht es in deiner Gesellschaft niemandem

We take this distribution as evidence that an empty operator without WH-features can cyclically move out of WH-islands, and that the restrictions derive from the illicit binding of topic phrases which cannot be picked up by such an operator. Relatively well-formed cases such as (36a) and the examples in (53) are then explained as follows: An empty operator moves cyclically through SpecCP, which is unfilled due to WH remaining in C-position. The trace of this operator will guarantee that the trace in IP is antecedent governed.

We must now explain why the trace of an empty operator, but not just any non-WH-trace can antecedent govern a lower trace. In (51) we have suggested that WH-movement will affect any WH' in C-position in such a way that both WH and WH' would have to appear in SpecCP in the syntax. Since the syntax (essentially S-structure) is more restrictive as far as this is concerned, we have said that such cases by disallowing antecedent government violate the ECP. At the same time it is common knowledge that LF processes are not subject to these severe restrictions. If they were, we could not derive the desired readings of multiple WH-questions. Let us now see why material, which can be linked to an empty operator is easier to licence in WH-islands than material that cannot be linked in such a way. Crucially, an empty operator is a "logical" element, i.e. something like WH, but without WH-features. If such an element is moved through SpecCP its trace will antecedent govern its extraction site. At the same time it will not affect the WH in C in the syntax, because it does not meet the input condition for convention (51). The question is now what will happen at LF where the WH in C has to be promoted to SpecCP as suggested by the condition on semantic well-formedness in (48). We propose that the situation is essentially the same as in other kinds of LF-movement, namely that 0 and WH can occupy one and the same SpecCP. Why should this be so with traces of 0-operators, but not with other (-WH) traces? The reason is likely to lie in the nature of the trace in question: Traces of operators share a [+operator]-feature with WH-elements, but other traces do not. Given that this is tenable, we arrive at the following derivation of an LF from S-structure:

- (58) SS: 0, ... [_{CP} e, [_{C'} WH, [... e_j ... e_i ...]]] ⇒
 LF: 0, ... [_{CP} e_i WH, [_{C'} [... e_j ... e_i ...]]]

Provided that other traces do not have the property of coexisting with WH in SpecCP, we can assume that a proper derivation fails by one of the following problems: Either WH cannot move to SpecCP, because SpecCP is already occupied by a trace with which WH cannot "combine", or the trace in SpecCP will be obliterated by WH, in which case the trace could not antecedent govern at LF. In both cases it is likely that an ECP-violation arises. According to Lasnik and Saito (1984) proper government is implemented as γ -marking, and

it is required that (with the exception of adjuncts) γ -marking applies at both S-structure and LF. In our case, γ -marking at LF by a trace of 0 would not be possible when the trace is obliterated by WH. On the other side, WH must occupy SpecCP at LF in order to γ -mark its dependent trace(s). Thus, both types of conflicts will ultimately result in ECP-violations.

Let us now compare the relatively well-formed example (36a) with the relatively ill-formed example (54a). These examples receive now the following LFs respectively:

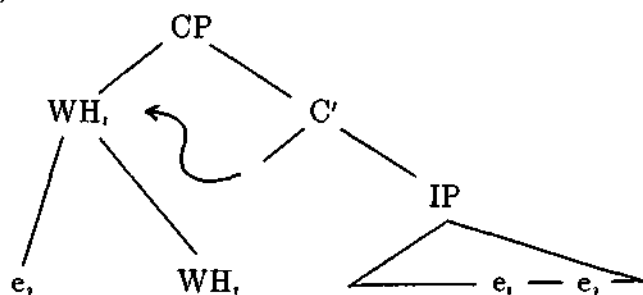
(59)a. ?Radios, 0, weiß ich nicht [_{CP=CP'} e_i WER_j [_{IP} e_j e_i repariert]]

b. ?*So schlampig, wüßte ich nicht [_{CP=CP'} WER_j [_{IP} e_j sein Auto e_i gewaschen hat]]

While by the mechanism of fusing the intermediate trace e_i with wer_j at LF some form of antecedent government can be obtained, this is not possible for elements that cannot be linked to a 0-operator. Thus, there is no way for antecedent government to arise in a representation like (59b), and an ECP-violation is predicted. The question is not so much whether some phrase XP can move in the disguise of a 0-operator, but whether there is an element in topic position which can identify such a 0-operator. For an ADV-phrase like *so schlampig* ('so messily') there seems to be no way of linking it to a 0-operator. For a bare plural NP like *Radios* it is obvious that it can be linked appropriately. We think that once this distinction is recognized it is futile to delve any deeper into the matter. The reason for this is that the intuitions about grammaticality often turn out to be unreliable, and many speakers may change their opinion from one example to the next. This may be interpreted as a measure of the speaker's ability to link a given phrase to a 0-operator.

The next question is why extractions from WH-island do still not yield perfect results. The reason undoubtedly is that antecedent government does not hold in the same way as it does in extractions from CPs whose head is a lexical complementizer. In the latter case antecedent government holds at both S-structure and LF in the same way. In the former it holds at S-structure, but only in a special way at LF, i.e. from a SpecCP-position in which both the WH-element and the trace of an empty operator occur. Let us assume that WH moves from C to SpecCP and that the trace in SpecCP has to adjoin to WH. This yields a structure as in (60):

(60)



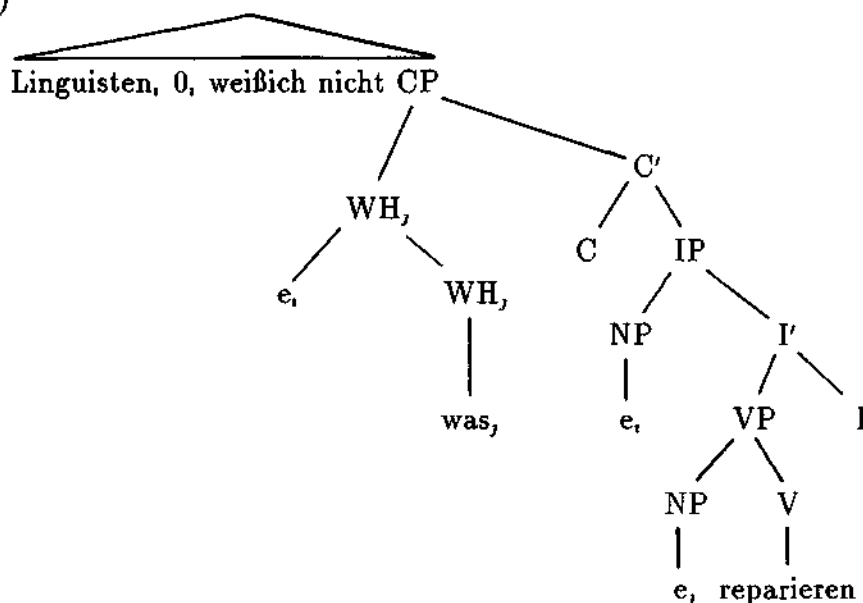
The intermediate trace e_j now appears in an operator position, but it does not antecedent govern the trace e_j inside IP in the strict sense because it fails to c-command it. This may be the correct explanation for the marginality of WH-island violations. Let us say

that in a situation like (60) antecedent government is “weak”, because it does not have the c-command property at LF. In order to make up for this deficit it is usually assumed that lexical government can rescue a construction. As we have argued, however, lexical government seems to have a different status in German than it does in English. We will turn to the status of lexical government in the next section. For the time being, it is necessary to point out an important empirical generalization, namely that the equivalent of lexical government in German seems to be that the trace be included in the VP. We have seen examples of WH-islands violations in which [-WH]-phrases were extracted which undoubtedly had the status of an adjunct and still the result was not as ungrammatical as one could expect, (e.g. (38a), (53a), (53b)). In German, these adjuncts always arise from inside a VP. As a contrast, extraction of the VP-external subject yields an ungrammatical result. This is shown by example (36b) which we reproduce here for convenience:

- (36)b. *Linguisten, weiß ich nicht [was, e, e, reparieren]
 linguists know I not what repair
 “As for linguists, I don’t know what (they) repair”

Given the trace of a 0-operator to WH, the LF-representation of (36b) will be:

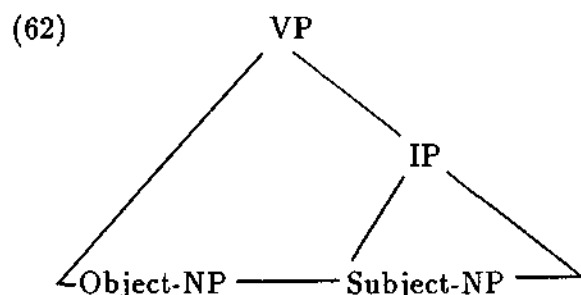
(61)



According to what we had said before, antecedent government is “weak” here, because the intermediate trace fails to c-command the extraction site. On the other hand, the 0-operator is not extracted from inside VP, but from a VP-external subject position. Let us assume that antecedent government of the kind we see in (61) is not enough to license a trace, but that in addition the trace has to be inside a VP.

This suggestion gives rise to a prediction for the extractability of subject-NPs in German. As is widely known, German is a scrambling language in which object-NPs, especially clitics and definite NPs, can be scrambled over the subject NP (see Fanselow, 1987; 1989,

Webelhuth, 1988, von Stechow and Sternfeld 1988: ch.12). Another way of seeing scrambling in German, and actually the one which I prefer, is to assume that under special conditions the subject NP can be generated inside VP, i.e. in close proximity to the inflected verb. In this case, the IP is as in (62) intertwined with the VP in such a way that the “span” of VP properly includes the “span” of IP (the notion of “span” is due to Hubert Haider):



If our generalization about extractability from VP is correct, then a “scrambled” structure should allow for less ungrammatical cases of WH-island violation by a subject-NP than a structure with an external subject-NP. This seems indeed a correct prediction, as the example in (63) show:

- (63) a. ?Linguisten, wüßte ich nicht [wie [dem Zirkusdirektor e, imponieren
 linguists know (subj.) I not how the circus-director impress
 könnten]]
 could
 “As for linguists, I would not know how they could impress the director of the
 circus.”
- b. ?Linguisten, wüßte ich nicht [was [uns e, zur Zeit Interessantes
 linguists know (subj.) I not what us(dat.) at-the-moment interesting
 mitzuteilen hätten]]
 to-tell have (subj.)
 “As for linguists, I would not know what interesting things they would have to tell
 us at the moment”

The case in (63a) is clear. *Imponieren* is an ergative verb in the sense of Burzio (1986). The standard assumption in the GB-framework is that the subject of such verbs arises VP-internally. In (63b) it is possible to argue for a VP-internal subject position, because in corresponding sentences such as *daß uns Linguisten zur Zeit nichts interessantes mitzuteilen haben* (“that linguists don’t have anything interesting to tell us at the moment”) the subject-NP is “closer” to the verb than the object-clitic. This clitic pronoun may be seen as extending the VP across IP, as shown in (62). When these assumptions are reasonable, the subject-NP would automatically arise inside the VP. As Kratzer (1989) has pointed out, the VP-internal versus VP-external status also shows up in so-called “split-NP constructions” (see v. Riemsdijk, (1989), Fanselow (1988), Bayer (1987)). In the following

sentences, the stranded prenominal quantifier *kein* is only grammatical when it is part of a VP-internal subject-NP:

- (64)a. ?Hochhäuser, weiß ich schon [warum [hier keine e, gebaut werden]]
 skyscrapers know I well why here none built are
 b. ?Hochhäuser, weiß ich schon [warum [hier keine e, eingestürzt sind]]
 skyscrapers know I well why here none collapsed are
 c. *Hochhäuser, weiß ich schon [warum [hier keine e, die Menschen froh stimmen]]
 skyscrapers know I well why here none the people happy make

For the passive, as in (64a), it can be assumed that the object-NP is not externalized and that nominative case is assigned VP-internally by INFL (see den Besten, 1985 and Haider, 1985). (64b) once again shows an ergative construction. According to Kratzer (1989), the complement of (64c) involves an individual level predicate in the sense of Carlson (1977). For reasons I cannot discuss here, such predicates require a VP-external subject-NP. Extraction from this VP-external position leads to firmly ungrammatical results.

Our suggestion that VP-internal status can make up for weak antecedent government at LF seems to be a proper generalization to capture both the facts about adjunct extraction and the facts of subject-extraction from WH-islands. The question is how we can implement antecedent government in such a way as to derive the proper distinction. When we say “antecedent government” we usually mean government from SpecCP. This is true for subject-traces which arise in SpecIP. For all the other traces that arise in VP, there is – according to Barriers – another antecedent governor, namely a trace which is left-adjoined to VP when move- α applies. Assuming adjunction to VP, the true picture shows the following difference:

- (65)a. **Extraction from VP-internal position**
 ... [_{CP} e_i'' WH [_{C'} [_{IP} ... [_{VP} e_i' [_{VP} ... e_i ...]]]]]]
 b. **Extraction from VP-external position**
 ... [_{CP} e_i' WH [_{C'} [_{IP} e_i [_{VP} ...]]]]

In (65a) the VP-internal trace is antecedent governed from the VP-adjoined position: the offending chain-link is (e_i'', e_i'); e_i is properly governed, but e_i' is at best “weakly” governed. This may account for the weak subjacency effect on otherwise acceptable WH-island violations in German. In (65b), however, e_i fails to be properly governed; there is no antecedent governor within CP and as a result the subject-trace e_i will not be antecedent governed in the regular sense at all. Examples such as those in (63) and (64a,b) conform to the structure in (65a), i.e. the subject-trace arises within the VP. Examples such as (36b) and (64c) conform to the structure in (65b), i.e. the only chain-link within CP that connects the trace in A-position is defective in the sense that it does not meet the c-command requirement of government and binding.

Although the difference between the different WH-island violations may be subtle, I do believe that it is real. Extraction of a [+WH]-phrase from a WH-island is sharply ungrammatical in German. No manipulation can rescue such a construction, and we are safe

in hypothesizing an ECP-violation. Subject-extraction of a VP-external [-WH]-phrase is – according to my own judgement – somewhat less offending, although not grammatical at all. In the proposed theory the difference between these two construction types is explained in the following way: WH-movement in the syntax forces any WH' in non-A position to SpecCP. Since at S-structure SpecCP cannot host two elements with the same features, one WH-item will be obliterated. Thus, one trace is not antecedent governed, and the structure is ruled out by the ECP. Crucially, VP-internality does not rescue such a construction, although there could be a VP-adjoined intermediate trace. Since, however, this intermediate trace fails to be governed from SpecCP at any level of representation, the situation is worse than in the case of the extraction of a VP-external [-WH]-item. In the latter case an empty operator leaves a trace in SpecCP at S-structure. This trace, we have assumed, will not be obliterated when WH moves from C to SpecCP in LF. At the same time, the trace can also not regularly govern the subject-trace, because it will have to adjoin to the moved WH-phrase. The third construction is the one in which a [-WH]-phrase or – in our terms – an empty operator corresponding to it is extracted from a VP-internal position. This is the least offending case. The result is almost grammatical. At S-structure there is a perfect A'-chain binding the position from which the 0-operator has been extracted. The only disturbing factor that intervenes is that WH-to-Spec CP at LF prevents the trace in SpecCP from truly antecedent governing the VP-adjoined trace. The 'squish' in grammaticality is summarized in the LF-representations of (66):

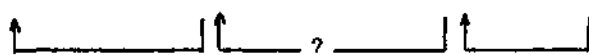
(66) a. *WH_i ... [_{CP} WH_i [_{IP} ... e_i ...]]



b. ?*O_i ... [_{CP} [_{WH} e_i WH] [_{IP} e_i [_{VP} ...]]]



c. ?O_i ... [_{CP} [_{WH} e_i WH] [_{IP} ... [_{VP} e_i [_{VP} ... e_i ...]]]]



As to my knowledge there is no work on German (or Dutch) which has tried to account for WH-violations of this sort. So the only comparison I can make at this point is the one with the general theory of grammar and the way of deriving ECP-effects.

Rizzi (1987), following a proposal in Stowell (1981), suggests to reorganize the ECP in such a way that its two halves have to be satisfied conjunctively, and not disjunctively as in Chomsky's classical version. SOV-languages such as German have properly head-governed subject-NPs, because I canonically governs to the left. Depending on what one has to say about the internal architecture of the "COMP-complex", there should be no subject/non-subject asymmetry in WH-island violation. Either they should all be ruled out, in case WH in SpecCP prevents intermediate traces, or they should all be grammatical, in case

SpecCP provides room for an intermediate trace. As we have seen, none of these logically possible predictions is fulfilled. What we see is rather a clear subject/object asymmetry for those cases in which the subject-NP is generated VP-externally. Clearly, proper government requires more than canonical government by I. There is no indication that I in German is sufficient. As can also be expected from this, German is not a null-subject language. My conclusion is that our implementation in terms of antecedent government is more suitable to derive the proper distinction.

What is the status of WH-island violations in English? For the core cases we have already discussed the relevant examples at the beginning of this section. Let us assume now that the mechanism of extraction from WH-islands is about the same in English and German. Given that lexical government is stronger in English than it is in German, we can derive the relevant contrast shown in (34) in our theory as well. What is new is that WH-elements in English embedded clauses will be moved to C, and only subsequently to SpecCP.

I would like to close this section with two pieces of evidence that this analysis is likely to be on the right track. The first piece of evidence is that English shows an asymmetry with respect to root-WH clauses and embedded WH-clauses. When non-subjects are moved by WH-movement, *do*-support applies in the former, but never in the latter:

- (67) a. What did John see?
- b. How did John fix the car?
- (68) a. *We don't know what did John see.
- b. *We don't know how did John fix the car.

An obvious way to prevent complement clauses of the kind in (68) is to assume that the C position is empty at D-structure and that it will be filled by the WH-element. My second piece of evidence is more delicate. It concerns the fact that not all WH-island violations are alike, even when adjunct phrases are moved. Consider the following contrast:

- (69) a. *On what did you wonder how Glenn Gould would have played Bach's
 'Italian Concerto'?
- b. ?On a cembalo I don't know how Glenn Gould would have played Bach's
 'Italian Concerto'?

According to the standard account of English WH-island violations it is mysterious how this fairly robust contrast could arise.¹⁵ Given that [_{PP} on NP] is an adjunct, both (69a) and (69b) should show an ECP-violation, because *how* would in either case prevent antecedent government. According to our own account the contrast will be derived in the same way as the corresponding contrast in German.

There is, however, one important difference between WH-island violations in German and English. Recall that while German in general disallows the movement of WH-elements from WH-islands, English does so only for WH-subjects and WH-adjuncts. Thus, an important difference between English and German seems to arise from the degree to which the

ECP can exploit lexical government (or θ -government). Although, we have concluded, there is no antecedent governor in any of the two language types when [+WH] moves out of a WH-island, lexically governed WH-material is far less offending in English than in German. In the next section we will try to derive this effect from properties which distinguish the two languages independently.

5. On the status of 'lexical government'

This section has a highly speculative character for reasons which we will see shortly. The question to be asked is where the difference comes from according to which lexically governed WH-arguments can extract in English but not in German. Let us first consider case morphology. An accusative object in English is to a large extent what an accusative object is in German. Of course, it is common knowledge that German has a more elaborate morphological case system than English, but this does not seem to lie at the heart of the difference. In fact, case doesn't seem to play a role at all. Consider the following English sentence (70a) and its ungrammatical German translation (70b):

- (70) a. ?Who did you ask how Bill could help?
b. *Wem hast du gefragt wie Bill helfen könnte?

While English *who* is three-ways ambiguous between nominative, accusative and dative case, German *wem* can only be a dative, and still the English sentence is tolerable while the German sentence is plainly ungrammatical. This strongly indicates that it cannot be the system of morphological case that should derive the difference. If this were true we would rather expect the opposite: According to certain assumptions about syntactic processing, (70a) involves a minor garden-path: *Who* will be associated with a potential gap after the verb *ask*, a step that has to be revised once the beginning of the complement clause becomes available. No such association can occur between the dative *wem* and the gap to the left of *gefragt*, because *fragen* does not assign dative case. If the problem would rest in the on-line resolvability of the argument structure, (70b) should be more acceptable to the speaker of German than (70a) to the speaker of English, clearly a wrong prediction. Let us therefore pursue another line of reasoning.

In what follows, we will digress from purely grammatical considerations and also include in our discussion considerations of syntactic processing. To begin, let us consider the compensatory "advantage" English has over German. Intuitively, the functional complement of the richer German case and agreement system is the rigid word order of English. Imagine extraction of an argument-NP in English. After encountering a displaced NP in operator position the processor can with high security hypothesize the gap which corresponds to this NP. This is very obviously not so in many sentences of German. Consider the following pair of WH-questions in English and German:

- (71) a. Who do you think that John has introduced to his wife yesterday at 5 o'clock?
b. Wen glaubst du daß Hans gestern um 5 Uhr seiner Frau vorgestellt hat?

The only place **who** can be associated with in (71a) is right after **introduced**. English, being a strict SVO-language, does not allow a trace to the left of the verb. On the other hand, it is well known (see Stowell, 1981) that an NP-object of the verb cannot travel away from the verb unless the special condition of Heavy-NP-Shift holds. In the German example (71b) the trace can be anywhere to the right of **Hans** and to the left of **vorgestellt**. This situation can directly be observed in the status of (im)possible echo-questions on both examples in (71):

- (72)a. You think that (*who) John (*who) has (*who) introduced (who) to his wife
 (*who) yesterday (*who) at 5 o'clock (*who)?
 b. Du glaubst, daß (??wen) Hans (wen) gestern (wen) um 5 Uhr (wen) seiner Frau
 (wen) vorgestellt (*wen) hat (*wen)?

In (72b) the various possible options exhibit different choices of scope, but what seems to be determined by S-structure constituent order here can in a case like (71b) only be determined by hypothesizing a trace at some point where the grammar allows for it. As we have seen, there is a variety of choices. The weakness of lexical government in German, as compared to its relative strength in English, could be derived on this basis from principles of processing efficiency with respect to filler-gaps dependencies. Such a principle could say that a trace which is not locally antecedent-governed can still be licit with respect to the ECP when – speaking now in processing terms – the gap can be identified by a unique location in the input. Such a claim, however, would have problems in dealing with the ungrammatical example (70b) above. In (70b) there is but a single position at which a trace can be licensed, namely to the left of **helfen**. All other logically possible choices would violate principles of German syntax. Obviously the ungrammaticality of (70b) cannot be directly connected to a processing problem.

If the weakness of lexical government is a direct reflex of the parser's being overwhelmed with too many choices for a single gap, then lexical government is predicted to hold in cases such as (70b), and our original attempt at an explanation of its ungrammaticality would be lost. Notice that the reverse situation also holds in English. There are well-known examples such as

- (73) a. What did you cook pea soup in?
 b. Which article did you read about?

Imagine on-line comprehension of these sentences. There is evidence from psycholinguistic research (see for instance Frazier (1987) on Dutch) that the WH-filler is "active" in the sense of its being inserted in the first place at which a gap can be hypothesized according to the grammar and presumably the preceding discourse/context. Intermediate representations will then arise which have to be revised later as more material of the clause becomes accessible. In (74) we indicate the time course of hypothesizing a gap by 1 for "earlier" and 2 for "finally".

- (74) a. What did you cook e_1 pea-soup in e_2 ?
 b. Which article did you read e_1 about e_2 ?

A transitional garden-path effect is predicted to arise when the e_1 -trace has to be rejected. The situation is, of course, not really comparable to the German example in (71/72b), because in (73/74) only one option will survive after the parse is completed, namely filling e_2 with the WH-filler. In German, the solution cannot be entirely decided by sentence grammar, because there is often more than one option for assigning a gap. What the examples in (73/74) do show, however, is that the processing complexity involved is unlikely to be amplified when the sentences in question appear as complements in WH-islands. The sentences in (75)

- (75) a. ?What did you ask how Bill would cook pea soup in?
 b. ?Which article did you ask how Bill read about?

remain about as grammatical or marginal as comparable sentences without the involvement of transitional garden paths.

As a consequence, a processing explanation of the strength/weakness of lexical government must not calculate a function of the number of possible traces in the actual input stream, because this number can vary quite drastically from one case to the next. What rather seems to be called for is an explanation which takes processing issues as "grammatized". This notion is not easy to define, but it should be clear what is meant by it intuitively. What we mean is that a processing strategy can become "hard-wired" in the computational system, such that the computational system employs this strategy even in cases where it is not called for functionally. It should be noticed that this is not out of the question, and that there are other domains in the syntax of Germanic which are likely to be the result of an adaption to processing difficulties.¹⁶ Speaking in purely representational terms, a trace could be almost anywhere between major constituents to the left of the verb as indicated in the German example (72b). Considering processing, it is quite impossible to see how the parser could determine a fixed locus for the trace in the input stream on-line.

A second and equally important point is that English has VO-order while German has OV-order, at least in the cases under discussion. Given that the verb is an extremely important indicator of the argument structure associated with it, it is plausible that once the processor has access to the verb it becomes easy to anticipate the argument structure. Actually, when we speak in processing terms, it is not only the argument structure as dictated by grammar, but the whole range of elements that may be associated as 'satellites' of the verb. It is clear that on the basis of the main verb, the processor can make immediate guesses about the associated material to come in English as well as in German V2-sentences in which the (inflected) main verb appears in second position. Crucially, however, such guesses are not reliable at all in the German cases under consideration. Here the verb is not available to the processor until the very end of the clause. It is not unlikely that in this situation the parser will wait until the end of the clause before final decisions about filler gap assignment will be made.¹⁷ Now it is easy to see in which way antecedent government can gain importance: Assume that normal processing of long distance dependencies goes step by step such that the processor begins from left to right at each CP to search for the gap corresponding to a

potential filler. In agreement with syntactic theory, we can assume that traces in SpecCP are the grammatical basis for this cyclic search. Under the assumption that syntactic perception is keyed to the CP-cycle, it can be imagined that any gap that appears within such a cycle can be identified from SpecCP. This is a rough sketch of the processing implementation of antecedent government. Assume now that due to a filled SpecCP-position the mechanics of the cyclic search for the gap breaks down. In this case it will be extremely important for the processor to precisely locate a potential gap inside the clause in which antecedent government failed. Of course, early access to a lexical governor (like V or P) with a missing complement can be imagined to help the processor out. This is roughly the situation we find in English. In German, however, the entire clause may have to be processed before anything conclusive can be said about its argument/adjunct-structure. Once the processor is prevented from postulating a trace in SpecCP, parsing the entire subjacent clause in search for a gap may turn out not to be sufficient for associating the filler with a gap. At least, such a gap could not be filled on-line, but only after the verb has been encountered.

To summarize, there are at least the following two obstacles: First, there is no upper limit as to where the trace could be hypothesized, because a potentially infinite number of adjunct-constituents can precede the lexical governor. Secondly, an object-trace (and in certain cases also a subject-trace) cannot be hypothesized before the verb is encountered. Due to the OV-order (which is obligatory when the C-position is filled with a complementizer or with a WH-item) a trace to the left of the verb can only be hypothesized after the verb becomes known. These two properties which distinguish German and Dutch from English are likely to create for the former two languages a massive problem for gap finding on the basis of lexical determination, which is the processing implementation of one half of the ECP. Seen from this perspective, certain WH-island violations in English seem to be tolerable as a result of lexical government, as GB-theory classically puts it. In any case, however, the two halves of the ECP cannot be of equal universality with respect to language use. We have shown that antecedent government always rescues a construction even when extraction out of a WH-island takes place. In those cases where antecedent government fails, lexical government can act only as a last-resort, and it can do so only when the processor can locate the lexically governed trace by virtue of a grammar-specific strategy. Evidence which supports the first claim is that in English WH-violations with long extraction of a WH-phrase are tolerable, but certainly not perfect. Evidence supporting the second claim is that they are tolerable in English while they are intolerable in German and Dutch.

Very clearly, these considerations must remain speculative as long as the relevant experimental evidence is missing. While there is evidence that in English traces in A-position invoke their antecedents in language comprehension (see Bever and McElree (1988), Nicol (1988), McElree and Bever (1989) and Nicol and Swinney (1989)), no work has as yet been done which addresses on-line processing of German or Dutch from a comparable perspective. One thing should, however, be clear even at this stage: The part of the ECP which says that a trace is properly governed when governed by a lexical governor has to meet with

the variability of lexical selection and with the fact that in certain languages the adjacency requirement is considerably relaxed or completely absent. Once antecedent government does not hold, these properties may have a profound influence on the licensing of a trace.

6. Summary and conclusion

In this article we have used the Vacuous Movement Hypothesis (VMH) in order to derive one major ECP-effect of English, namely the *that*-trace effect. Since German as well as Dutch are not subject to the VMH in the relevant cases, the absence of the effect in these languages is an immediate consequence of our approach. One reason for adopting the VMH to explain the *that*-trace effect was that the Minimality Condition (MC) of Chomsky (1986) as well as Rizzi's (1987) approach have problems in explaining the absence of tolerable WH-island violations in German. We could show that WH-island violations have to be differentiated as to the nature of the constituent to be extracted. For German and Dutch it turns out that extraction of WH-phrases is generally impossible, whereas it is well known that the extraction of θ -governed WH-objects leads to tolerable results in English. Our conclusion was that in both languages the WH-phrase in SpecCP obliterates the position through which WH-material has to move and that a trace in IP can only be legitimized by lexical government (or θ -government). The interesting point is in which way this obliteration occurs. We have shown that a number of facts suggests that WH-phrases in German arise in C and not in SpecCP. This would leave room for an intermediate trace in SpecCP. For long WH-movement we have assumed, however, that the WH-element in C-position is affected already in the syntax, and that as a consequence antecedent government never holds in these constructions. When [-WH]-material is extracted, the situation was found to be more complex. The proper generalization here seems to be that those VP-internal elements can extract which can be linked to an empty operator. Our conclusion was that an empty operator can move in the syntax through SpecCP leaving an intermediate trace. At LF, the WH-element in C moves to SpecCP and adjoins this trace. Since according to the Barriers theory movement from VP proceeds via adjunction to VP, the VP-internal trace will be antecedent governed. The intermediate trace in SpecCP, however, fails to c-command the VP-adjoined trace at LF. This was taken to be the reason for the reduced grammaticality of these still tolerable WH-island violations. Extraction of VP-external empty operators corresponding to the subject-NP fail to be antecedent governed in the normal sense, because at LF the (adjoined) trace in SpecCP does not c-command the extraction site. The remaining question was then why English can make up for a lack of antecedent government when lexical government holds, while German and Dutch cannot. It was suggested that there exists a processing difference between the two language types which makes the processor entirely dependent on antecedent government in the latter. Our core assumption was that it is mostly very easy to hypothesize a trace in English, due to the rigid word order and the head-first character of the language.¹⁸ In German it is not so obvious how a trace can be hypothesized at a unique point in the input stream. We argued that it may be possible in

certain cases, but it is very often the case that the precise location of the gap cannot be determined. Furthermore since German retains SOV-order in the relevant cases, a gap to the left of the verb could not be θ -marked before the verb is found in left-to-right processing.¹⁹

Given this analysis is tenable, it has important consequences for the nature of the ECP. According to the classical approach proposed in Chomsky (1981) and assumed in much subsequent work, the two clauses of the ECP may apply disjunctively. There are other proposals according to which they have to apply conjunctively. (See Rizzi, 1987:section 2 and the references given therein). A recent conjunctive version of the ECP is given in (76):

(76) Conjunctive ECP (Rizzi, 1987)

A non-pronominal empty category is

- (i) canonically head-governed, and
- (ii) antecedent governed or θ -governed

It is obvious that the classical formulation does not account for the facts of German. The reformulation in (76) does not account either for the complex array of facts that make German and English both similar and different. First of all, there is no reason why WH-islands violation should be less tolerable in German than in English. If one could argue that antecedent government has an effect in such cases, both languages should allow for the extraction of θ -governed elements. As a matter of fact, however, German does so only when [-WH]-items get extracted. If antecedent government is impossible, both languages should exhibit clear ECP-violations, which they don't. Problems also arise in connection with *that*-trace violations. In German, an SOV-language, the *that*-trace effect is (correctly) predicted by (76) to be absent, because antecedent government is fulfilled as well as canonical head-government by I. Given that in the extraction of [-WH]-phrases out of WH-islands antecedent government is guaranteed, subject-NPs (or empty operators in subject-position) should also extract. This again is contrary to the facts. Our own proposal has to acknowledge the fact that the ECP has a strong part which is probably universal and a weak part which is subject to parametric variation which may ultimately derive from the choice of direction of government. Furthermore we have to acknowledge Rizzi's (1987) observation that an unqualified notion of θ -government is not able to rule out undesirable cases such as the long movement of semantically selected adverbs. Let us therefore tentatively claim that θ -government is to be understood in the sense of Rizzi (1989) as "assigning a referential index". (See the appendix for further discussion). Taking these considerations into account, we propose the version of the ECP given in (77).

(77) Parameterized ECP

A non-pronominal empty category is properly governed iff

- (i) it is antecedent governed or
- (ii) θ -government identifies to the right of the governor a unique location of the empty category

Full grammaticality arises only when the strong part (77i) is fulfilled. In this case, θ -government becomes obsolete. Semi-grammaticality arises when only the weak part (77ii) is fulfilled. Our hypothesis is that this case can only arise in languages, which generally guarantee the precise localization of a θ -governed gap in the input stream. In those languages in which θ -government can identify (on-line) a gap the lack of antecedent government will lead to a weaker violation than an ECP-violation, namely to a subjacency violation. In languages in which (on-line) gap identification via θ -government is potentially inhibited as a result of the leftward orientation of governors, antecedent government becomes crucial. Once antecedent government does not hold (in any form), the violation is predicted to be stronger than a mere subjacency violation. According to our account, it would be an ECP-violation.

Appendix: On referential indices

What has been referred to in this article as Rizzi (1987) has undergone revisions which touch on some of the issues treated here. In Rizzi (1989) the conjunctive ECP of Rizzi (1987) is revised in such a way that the notion of θ -government disappears completely. (See also Cinque, 1989). Before we elaborate on this, however, yet another revision should be mentioned, namely the one that replaces the notion of "canonical head government".

In Rizzi (1989) the notion of (linear) "canonical head government" is given up in favor of a non-linear hierarchical definition of head government. Proper head government now requires that YP is governed in the immediate projection of X^o . This step makes a crucial prediction for German, because now the pre-VP subject position is not head-governed by I(NFL). The immediate projection of I contains at best the VP, but not the external argument. Since the requirement of head government is not satisfied, German and Dutch are predicted to show the *that*-trace effect. The evidence in favor of this prediction is meager. (See however Fanselow, 1987 for some evidence).²⁰

Let us now turn to the more central issue of the abandonment of θ -government. As we said already, Rizzi made the important observation that non-NP categories which are semantically selected (and thus θ -governed) cannot be extracted from WH-islands. For instance, although the verb *weigh* can select both a patient-NP like *apples* or a measure phrase like *200 lbs*, only the former is available in a case like (1):

- (1) ?What_i did John wonder how to weigh *e_i*?

To answer the question stated in (1) the answer *apples* would be appropriate, but not *200 lbs*. Rizzi's proposal is then that only those categories can circumvent the requirement of antecedent government which carry a 'referential index'. A referential index is assigned to those XPs (mainly NPs, but also certain PPs) which belong to the set {agent, patient, theme, ...} of θ -theory. Other elements such as manner adverbs, measure phrases, idiom chunks etc. do not obtain such an index. This leads to a crucial difference in A'-dependencies: R-indexed A'-chains and non-R-indexed A'-chains. The difference is expressed in the pair of examples in (2):

- (2) a. Who_k did you see e_k?
 b. How did you behave e?

This difference is then exploited in the assumption that there is an A'-binding relation that depends on an R-index, but is free of the locality constraints which hold for antecedent government.

- (3) X binds Y iff
 (i) X c-commands Y and
 (ii) X and Y have the same referential index

According to (3), a non-pronominal empty category must be properly head governed, but can be either bound at long distance via a shared referential index or antecedent-governed via a chain of locally accessible traces. This theory makes a wide range of correct predictions, also cross-linguistically. It is immediately clear that the problem of s-selected non-referential categories disappears. They simply do not have a referential index. Once their traces cannot be identified via antecedent government, they violate the ECP. The measure-reading on (1) amounts to such a violation. It is easy to see how the theory predicts the usual argument/adjunct asymmetries. Subject extraction from *that*-clauses or WH-islands would be licit by virtue of R-binding, but it does not fulfill the requirement that the trace be properly head governed: The VP-external subject trace is not in the immediate projection of a head-governor, here I(NFL).

As Rizzi says, subject extraction from WH-islands is marginally possible in Italian. One of his examples is shown in (4):

- (4) ?Che studente, non sai come, potrai risolvere il problema e, e_j?
 Which student not (you)know how could solve the problem
 'Which student don't you know how could solve the problem?'

According to Rizzi (1982), the subject-NP can be postposed in Italian, thus getting into the government domain of the verb.²¹ In this case, the subject trace is head-governed, and long subject extraction is predicted to be licensed by both R-binding and proper head government.

In evaluating Rizzi's (1989) theory as well as Cinque's (1989) proposal for refining the theory of referentiality I see a strong convergence of ideas concerning the mechanism which underlies acceptable WH-island violations. Rizzi as well as Cinque argue in favor of locally unconstrained A'-binding for referentially indexed phrases. In this article I have independently proposed a mechanism which allows "referential" XPs which are [-WH] to antecedent-govern a trace in WH-islands at S-structure. According to this latter analysis antecedent government (at least in some special form) will still hold in such a case. What is crucial, however, is that in both approaches "referentiality" must somehow be met.

In the rest of this appendix I would like to point out four problems which I still see in connection with Rizzi's (1989) theory. Three of them arise from German examples which were discussed in this article or closely related examples.

1. Subjects in VP

German not being a null-subject language does not have the option of subject postposing that Italian has. Still, German has other means of having a subject NP close to the verb. Following Burzio (1986), one can argue that ergative subjects arise inside VP (see den Besten, 1985 and Grewendorf, 1986). Arguably this makes them head-governed by V. Since at least WH-NPs with a restricted quantifier (see Pesetsky, 1986, Dobrovie-Sorin, 1988 and Cinque, 1989) can be considered as “referential” or “discourse-linked”, such WH-NPs should be able to extract from WH-island without causing an ECP-violation. As the following example shows, however, the result is still ungrammatical enough to argue in favor of an ECP-violation:

- (5) *[Welcher Fehler], wusste er nicht [wem [e, unterlaufen ist]]?
which mistake knew he not whom happened is
‘Of which mistake did he not know to whom it happened?’

2. WH-Objects

Object-extraction does not lead to any better results in German. Whatever the degree of discourse-linking/referentiality is, the following example shows that something else must rule out (6):

- (6) *[Welches der Kinder], wußte der Vater nicht wann er e, abholen sollte
which of the children knew the father not when he pick-up should
“Which of the children did father not know when to pick up?”

Rizzi’s theory would again predict that WH-phrases with a “referential” index can be extracted from WH-islands. They are head governed, and they are of the adequate semantic type, but the result is ungrammatical as ever. There is thus no reason to pay any more attention to the status of subject-NPs (although this is relevant elsewhere, as we have seen in connection with [-WH]-NPs). The problem seems to rest entirely in the [+WH]-status of the category to be extracted.

3. An unexpected ECP-effect in Bavarian

Rizzi as well as Cinque argue for non-local A'-binding in the sense of (3) above. We have already seen that such an account does not automatically cover all the facts of German WH-island violations. Something in addition must certainly be said about the long extraction of WH-phrases. Let us now see what non-local binding predicts for those cases which were more or less unproblematic in German, namely those in which a [-WH]-phrase seems to be extracted from a WH-clause. Recall that we have argued at length that these cases exist because the WH-phrase of the embedded clause occupies the C-position at S-structure, leaving the possibility that a null-operator can extract via SpecCP. This amounts to a form

of antecedent government. The Rizzi/Cinque theory, on the other hand, is designed in such a way that R-binding makes antecedent government superfluous. Let us test this prediction.

Bavarian allows for a simultaneous filling of C and SpecCP with lexical material. Imagine that the C-position is filled with *daß* and SpecCP is filled with some WH-phrase. According to the Rizzi/Cinque account this situation should still allow for non-local R-binding. According to the antecedent government account, however, it should lead to a severe degradation of grammaticality. The following examples (for which only the syntax is Bavarian) indicate that the latter prediction is the correct one:

- (7)a. *Radios, weiß ich nicht [wer, [daß [e, e, repariert]]]
 “Radios, I don’t know who that repairs”
 b. *Radios, wollte niemand wissen [wie, [daß [wir e, e, reparieren]]]
 “Radios, nobody wanted to know how that we repair”
 c. *Radios, ist im Moment unklar [wann, [daß [wir e, e, wieder reparieren]]]
 “Radios, it is at the moment unclear when that we repair again”

The examples in (7) have a “doubly filled COMP”, thus precluding material to move through SpecCP. On the other hand, when the WH-element is in C-position, an empty operator can use SpecCP as an escape position. In this case there is only a weak subjacency violation due to the fact that WH must move to SpecCP at LF (see section 4.2 above).

4. VP-movement in Italian

Referential indices are only assigned to elements which “refer” according to θ -theory and semantic constraints on referentiality. Elements that do not bear a θ -role, e.g. adjuncts, do not qualify. Consider now VPs. VPs certainly do not bear a θ -role, given the plausible assumption that something that assigns a θ -role does not in turn bear one. This rules out VPs as carriers of an R-index. The following Italian example (provided by Lauredanna Frauenfelder, pers. comm.) disconfirms this prediction:

- (8) ?[Andare a casa], mi ciedi [come [Gianni potra e,]]
 go home me ask how Gianni can (fut)
 “Go home, I ask myself how Gianni will be able to”

I must admit that there are speakers of Italian who do not accept (8) or similar sentences. One would have to see whether there are dialectal differences.

5. Consequences

What are the possible conclusions to be drawn from these observations? One is that WH-phrases may after all not fall under the [+/-referential] distinction. Whatever the proper formulation of the theory will be, the overall success of this distinction (see in particular Cinque, 1989) should preclude this conclusion as premature.

A more promising target of attack is the suggestion that referential XPs can bind a trace at long distance. Given standard assumptions about binding (even A'-binding), this suggestion is indeed unusual. As a matter of fact, discourse binding in the languages under consideration is always binding of a lexical pronoun, not the binding of an empty category.²¹ This can be seen in the following example:

- (9) [A woman], walked in. John kissed her,/*e.

Under our more conservative proposal, long-distance binding of a trace is not necessary. The distinctions which are required can be captured by independently motivated assumptions about the architecture of C and SpecCP. According to this analysis, WH-phrases moving

out of WH-islands violate antecedent government on one of the WH-traces involved. Thus, an identification system independent of antecedent government has to be invoked in order to let such structures pass the ECP. In the present article (see (77)) we have proposed that θ -government has to be able to identify a unique location of the trace in the input. The trace and its antecedent are coindexed (due to move- α), but when antecedent government is not available, the identification of the trace has to rely on θ -government. Given Rizzi's observations about the failure of θ -government for s-selected but non-referential XPs, this requirement seems too general. As it is formulated, it would still permit measure phrases and the like to circumvent antecedent government. It is, however, possible to further qualify the notion of θ -government by relaxing its identification with s-selection. In this article we have proposed that null-operators with a trace in a WH-clause can only be licensed by a topic phrase which is coindexed with it. A topic-VP like the one in (9) could also be coindexed in this way. At the same time this proposal rules out the cases which are also ruled out by the Rizzi/Cinque account. This is so, because there is no way of relating certain elements (including quantified NPs) to null-operators/pronominals.

The ungrammatical Bavarian examples in (7) (as compared with their grammatical counterparts in section 4 of this article) indicate that long-distance binding is too strong a machinery, and that antecedent government in one form or the other must be involved here.

More difficult to account for is the situation in which WH-elements move out of WH-clauses. As we have seen, these cases are mostly (and in German always) worse than those in which a [-WH]-element moves "long distance". It is only for this case that we have proposed a complete failure of antecedent government, and it is only for this case that we invoke an identification system independent of antecedent government. Seen from a processing point of view, one can imagine that in the absence of an antecedent governor the processor is unable to establish traces for certain XPs, as the current theory of Rizzi predicts. This, however, does not necessitate the conclusion that there is non-local A'-binding. In my opinion it rather requires the language to be such that it allows for the establishment of a trace on the basis of a lexical trigger alone. This comes close to the classical disjunctive version of the ECP, but the value one will now ascribe to its second half is less strong than in Chomsky's original version. Under this somewhat speculative assumption it could be possible to not only account for the English and Italian data, but also for the intricate pattern we see in Dutch and German.

Notes

1. The relevant definitions in *Barriers* consider only X^o -or X^{max} -categories, but no intermediate categories such as C' .
2. Intuitively this seems unattractive, because minimality is usually taken to refer to cases in which both the distant and the close governor have the potential to govern an empty category, e.g. in $[_{VP} V [_{PP} P \text{ trace}]]$ P is the governor of the trace, but not V . English complementizers, on the other hand, are certainly not governors in this sense. If they were, that-trace effects would not arise at all.
3. Rizzi, of course, refers in this connection to the role of the overt complementizer in creating a minimality barrier. In chapter 8 of the *Barriers*-text, however, it is unclear how the VP could be crossed by an adjunct, given that the strong version of the MC is adopted. The reason is the following: V is a (lexical) head whose projection immediately dominates V . The adjunct is presumably included in and not adjoined to the VP. (see *Barriers*, p. 20). Thus, unless a special amendment of the notion "immediate dominance" is provided which says that adjoined positions count as immediately dominated, VP will be a minimality barrier. An amendment along this line, however, seems out of the question, because positions adjoined to XP are taken not to be dominated by XP. (see definition (12), p.7 in *Barriers*). Once adjoined positions are not dominated, they can certainly not be immediately dominated. If this is a correct interpretation of the strong version of the MC, not only that in (5), but also the verb *fix* would create a minimality barrier.
4. See George (1980), where this version is introduced as the "Strong Vacuous Application Prohibition". George distinguishes it from the "Weak Vacuous Application Prohibition", which he ascribes to Chomsky (1973). According to the latter, vacuously applying restructuring rules are confined to applying only in the readjustment component of the grammar.
5. The same arguments hold for cases of LF-moved adjuncts such as (106b) of *Barriers* *Who knows what John did how. An open question which I will not pursue here is why superiority effects are absent in German. (See Haider, 1983 who assumes that the subject-position in German is properly governed by the verb. We will see later that this would not make the correct predictions for German WH-island violations).
6. One could imagine that do-support with stress on do can apply to cases of long subject-WH-movement. In this case, do-support would signal overt movement of a category. Indeed, such examples are grammatical, as (i) shows:
 (i) Who do you think DOES understand this theory?
 Assuming that the auxiliary occupies the same position as the complementizer that otherwise, (i) would have the representation in (ii):
 (ii) Who, do you think [_{CPe}, [_CDOES [_{IPe}, [understand this theory]]]]
 Angelika Kratzer (p.c.) pointed out to me, however, that (i) is unlikely to involve what is at present referred to by "do-support", namely movement of do from I to C.

According to Kratzer, (i) is a case of insertion of **do** in I (such that the inflectional morpheme can be spelled out without the verb moving to I). That she is right about this, is shown by the fact that stressed **do** and the complementizer **that** are not in complementary distribution:

(iii) (context: Mary doesn't buy books when she goes to London. So...)

what do you think that Mary **DOES** buy when she goes to London?

If **DOES** would be in C-position in (iii), **that** could not be there as well, given the present assumptions about I-to-C-movement. We conclude that cases such as (i) do not constitute any evidence in favor the VMH.

7. The \-slash should indicate that the category on the left side of \ is mapped onto the category on the right side.

My adoption of Haider's idea of matching projections does not mean that I would subscribe to all corollaries of his theory. For instance, he argues that English declarative sentences "imitate" the CP-structure by collapsing the CP into IP in the sense of matching projections. This will yield the same structure for both declaratives and subject-moved WH questions:

(i) [_{CP} \ IP he [_C \ I' will [_I P believe that]]]

(ii) [_{CP} \ IP who [_C \ I' will [_I P believe that]]]

One problem with this is that adverb placement is more restricted in (ii) than it is in (i). A sentential adverb like **probably** can exceptionally appear before the I-position, but certainly not before the C-position. As (iii) and (iv) show, the examples in (i) and (ii) reflect this distinction:

(iii) He **probably** will believe that

(iv) *Who **probably** will believe that?

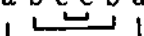
Given that IP and CP are collapsed in all those cases in which a subject-moved WH-clause resembles an IP, it is unclear how this distinction could be expressed. In my view, the two examples are nothing but instances of IP and CP respectively. This assigns a much narrower range of application to matching projections than Haider intends.

8. Haider himself proposes something different here, namely that a matching projection involving a CP is open for government by the matrix verb. One of his examples (quoted from Jespersen) is **Children whom we thought need care**. Whatever the status of this example is, it certainly does not reflect present day usage, as one can see in comparable ungrammatical cases such as ***We thought them (to) need care**.

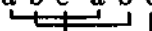
Building an explanation of the **that**-trace effect of lexical government, Haider suggests that the subject-trace in a headless CP will be governed by the matrix-verb, whereas it cannot similarly be governed when the CP has an overt head. Given the unclear status of the concept "lexical government into CP", it seems to us more promising to develop an ECP-explanation along the lines of antecedent government, as suggested in this paper.

9. Strictly speaking, antecedent government will also hold in the adjunct cases if the Barriers-framework is assumed. There is evidence from English (see Chomsky, 1986) and absolute certainty in German that how-adjuncts arise inside VP. By adjunction to VP the adjunct-trace will be antecedent governed. In any case, however, the A'-chain will be interrupted by CP which is a barrier. If nothing else is said, this is what we mean by 'violation of antecedent government' for the rest of this paper. We will turn to a special case towards the end of section 4.
10. For German see Cardinaletti (1986). Following Koster (1978), Cardinaletti proposes to base-generate a preverbal phrase in German root-clauses in a separate TOP-position (see Chomsky, 1977). This TOP-position is then coindexed with SpecCP. SpecCP may contain WH- and d-words or an empty operator. Cardinaletti's reason to do this is based on Cinque's (1985) observation that syntactic variables can only be identified by operators. Thus, a non-operator phrase in SpecCP could not license a variable. With the mechanism of a coindexation between TOP and an empty operator, a variable can be licensed. Notice, however, that once it is assumed that the pre-verbal position is always an operator position binding a variable, this approach faces problems e.g. with sentential adverbs. They can appear pre-verbally, but they do not correspond to a d-pronominal. A more serious problem seems to be that Cardinaletti follows Cinque (1984) in adopting *pro* as the empty operator. A cornerstone of Cinque's theory, however, is that *pro* can only correspond to categories which bear the same syntactic features, i.e. to NPs only. I will not pursue this more general solution, but rather reserve the Cardinaletti-Cinque-Koster proposal as merely an option of UG. It is, however, interesting to see that categories which are problematic for Cardinaletti's theory are also misbehaved when moved out of WH-islands. The b.-examples below, which are derived from the (grammatical) a.-examples, demonstrate this:
- (i) a. Ich weiß nicht wer sofort zu spielen anfangen kann
I know not who immediately to play begin can
b. *Sofort, weiß ich nicht [wer e_i zu spielen anfangen kann]
- (ii) a. Ich weiß nicht wer jede dieser Sonaten spielen kann
I know not who each (of) these sonates play can
b. *Jede dieser Sonaten_i weiß ich nicht [wer e_i spielen kann]
- Whatever the precise restriction at work is, the ill-formedness of the b.-examples corresponds to the well-known ill-formedness of topic constructions such as
- (iii) a. *Sofort, das kann er nicht zu spielen anfangen
immediately this can he not to play begin
b. *Jede dieser Sonaten, die kann er nicht spielen
each (of) these sonatas them can he not play
11. For further details see Chomsky (1986: ch. 11). For attempts to derive the HMC from more general principles see Pollock (1989) and Chomsky (1989). I will refrain from going into these details.

12. See also Haider (1988, note 4) where the same idea is adopted.
13. If this is to be seen as movement from C to SpecCP our concerns about (45b) above would return. Let us therefore assume that this movement is not of the trace-theoretic kind. To put it another way, by moving WH from C to the adjacent SpecCP C' and CP will collapse.
14. Kayne (1989) assumes clitics in Romance to be able to move via I through C to the I of the embedding clause. As to my knowledge there is, however, never an instance where this kind of movement can proceed any further, say, another clause up. And certainly there are theoretical alternatives to Kayne's approach which would not involve movement through C at all.
15. Although we cannot go into a discussion of Scandinavian, it is interesting to see that virtually all the standard examples for WH-island violations which are found in the contributions in Engdahl and Ejerhed (1982) deal with [-WH] topic phrases. According to Suzanne Schlyter (pers. comm.) these cases are easily acceptable, whereas those involving long movement of [+WH]-phrases are felt to be problematic by many speakers.
16. One domain in which empirical exploration has begun is the syntax of the verbal complex in Germanic SOV-language. It is a well-known fact that German and Dutch as well as their respective dialects organize the verb-cluster differently. Textbook knowledge tells us that German retains a strictly nested order between arguments and verbs, i.e.

a b c c b a


while Dutch has shifted to a crossed order, i.e.

a b c a b c


In fact, however, the situation in German is far more complicated, and it is clear that given a certain level of complexity the dependencies begin to resemble Dutch. To my knowledge there is nothing in contemporary theory of linguistic representations that could explain this variability. There is, however, a way to look at these intricacies from a processing point of view. See Bach, Brown and Marslen-Wilson (1986). But certainly any processing explanation of these facts has to acknowledge that the order of verbs in Dutch has become "hard-wired" in the sense that even sentences which are easy to process in German, i.e. with nested dependencies, are often felt to be ungrammatical by Dutch speakers when presented in nested order. For example

(i) weil wir den Mann singen hörten (German)

(ii) a. omdat wij de man hoorden zingen (Dutch)

b. ?*omdat wij de man zingen hoorden

‘because we heard the man sing’

Insightful remarks on this are found as early as Evers (1975) for Dutch and Standard-German and Lötscher (1978) for Swiss-German, but until now linguistic theorizing has not much cared about these lines of reasoning.

17. The qualification “final” is important here, because there is experimental evidence that the processor makes immediate decisions which later on may have to be revised. Frazier (1987) investigated the processing of ambiguous and non-ambiguous Dutch relative clauses in which either a subject- or an object-gap had to be filled with the relative pronoun. For the ambiguous clauses she found a clear subject preference (74%), while the reading times for unambiguous relatives with subject gaps were only non-significantly faster than those for sentences with object gaps. In addition, more errors were made on (unambiguous) object-gaps than on subject-gaps. These results support the conclusion that the processor does make certain commitments before the entire clause is parsed. The situation described in Frazier’s experiment, however, is the normal situation in which an antecedent in SpecCP initiates the search for a gap in IP. It does not necessarily bear on the issue referred to in this article, because pre-VP subject gaps may not be expected at all in WH-islands. Thus, one would have to find out at which point in time the listener makes a decision about an object- (or VP-internal subject-)gap. What is needed is a more direct window into the search for a gap, as in the antecedent-priming study by Nicol (1988).

18. We say “mostly” because there are well-known problems with datives in English. Fodor (1978) provides the following pairs:

(i) a. What_i did John give Mary e_i?

b. ?*Who_i did John give e_i a book?

(ii) a. The patient whom_i the nurse brought the doctor e_i hated rainy days.

b. ?*The patient whom_i the nurse brought e_i the doctor hated rainy days.

In the b.-sentences the parser is likely to skip the trace and assign the dative to the first lexical NP which is encountered. Following an earlier proposal by Jackendoff and Culicover (1971), Fodor suggests a constraint which says that extraction of a category X is blocked when the trace of X is followed by a category of the same type, and both categories X are contiguous in D-structure (see Fodor, 1978:444). See also the discussion in Hankamer (1973) and related literature quoted in Fodor (1978). This constraint, known as the ‘XX-Extraction Constraint’ does not apply to cases such as (iii) (=Fodor’s (17))

- (iii) Who_i did Tom ask (*e_i) Meg to persuade (*e_i) Jill to inform (*e_i) Ted that Bob has spoken to e_i?

because the parser will correctly take the lexical NPs following each verb as the argument of the verb. As far as I can see, the XX-Extraction Constraint also does not clash with the Active-Filler Hypothesis of Frazier (1987), which indeed predicts minor garden path effects in a case like (iii). These effects are at any rate quite tiny, and the parser will easily recover from a misanalysis by encountering the immediately following NP.

19. On the basis of this argument, one could assume that things improve when the extraction site is to the right of the verb. There is, however, good evidence from both Dutch, German and other SOV-languages that no extraction can take place from extraposed material unless the extraposed material is itself an argument-CP, which as such may allow for the independent mechanism of COMP-to-COMP movement. Koster (1986:ch.4) observes that reorderings which deviate from the canonical order lead the reordered material to become absolute islands for extraction. Pronominal adverbs can split the pronominal from a postposition like *daarmee* in Dutch and *damit* in German when this postpositional PP is to the left of V. As (i) and (ii) show, this PP can be extraposed.

- (i) omdat hij niet heeft gespeeld daarmee
since he not has played it-with

- (ii) weil er nicht gespielt hat damit

But then P-stranding is blocked:

- (iii) *Daar, heeft hij niet gespeeld [e, mee]

- (iv) *Da, hat er nicht gespielt [e, mit]

this has he not played with

‘He hasn’t played with this’

Parallel effects were demonstrated for quantifier raising in LF for German by Bayer (1989) and for WH-in-situ in Bengali by Bayer and Dasgupta (in preparation). Assuming that the processor obeys grammatical constraints, such structures which at first sight could be assumed to improve parsability turn out to be irrelevant, because parsability does not make ungrammatical sentences grammatical.

20. Notice also that the present definition of proper head government would not allow the extraction of a dative PP in English, if the PP is not in the immediate projection of V, i.e. if we do not assume something like ternary branching as in [₁ P V NP PP].
21. See Heim (1982) for extensive discussion of pronouns as bound variables. Rizzi (1986: section 2) shows that certain null-objects (in Italian) cannot be variables, but must be of the category ‘pro’. In the following example

- (i) La buona musica riconcilia [pro] con se stessi

‘Good music reconciles with oneself’

the empty category receives an arbitrary interpretation, i.e. an interpretation that is generally not available for variables.

22. It is not really clear how the subject-trace can be head-governed by the verb **risolvere**, which governs the NP **il problema**, but for the sake of discussion let us assume that this problem can be solved.

References

- Bach, E., C. Brown and W. Marslen-Wilson (1986), Crossed and nested dependencies in German and Dutch: A psycholinguistic study. *Language and Cognitive Processes* 1. 249-262.
- Baker, C.L. (1970), Notes on the description of English questions: the role of an abstract question morpheme. *Foundations of Language* 6. 197-219.
- Bayer, J. (1984), COMP in Bavarian syntax. *The Linguistic Review* 3, 209-274.
- Bayer, J. (1987), The syntax of scalar particles and so-called 'floating quantifiers'. Ms. Max-Planck-Institute for Psycholinguistics, Nijmegen.
- Bayer, J. (1989), Interpretive islands: evidence for connectedness and global harmony in logical form. in: Grewendorf, G. and W. Sternefeld (eds.),
- Bayer, J. and P. Dasgupta (in preparation), Complementation, orientation and WH-in-situ in Bengali. Ms. Max-Planck-Institute for Psycholinguistics, Nijmegen.
- den Besten, H. (1977/1983), On the interaction of root transformations and lexical deletive rules. Ms., published 1983 in: Abraham, W. (ed.), *On the Formal Syntax of the Westgermania*. Amsterdam: John Benjamins.
- den Besten, H. (1985), The ergative hypothesis and free word order in Dutch and German. in: Toman, J. (ed.), *Studies in German Grammar*. Dordrecht: Foris.
- Bever, T. and B. McElree (1988), Empty categories access their antecedents during comprehension. *Linguistic Inquiry* 19. 35-43.
- Burzio, L. (1986), *Italian Syntax. A Government-Binding Approach*. Dordrecht: Reidel.
- Cardinaletti, A. (1986), Topicalization in German: Movement to COMP or base-generation in TOP? *Groninger Arbeiten zur Germanistischen Linguistik* 28. 202-231.
- Carlson, G. (1977), *Reference to Kinds in English*. Doctoral dissertation, University of Massachusetts, Amherst.
- Chomsky, N. (1973), Conditions on transformations. in: Anderson, S. and P. Kiparsky (eds.), *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston.
- Chomsky, N. (1977), On WH-movement. in Culicover, P., T. Wasow and A. Akmajian (eds.), *Formal Syntax*. New York. Academic Press. 71-132.
- Chomsky, N. (1981), *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. (1986), *Barriers*. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (1989), Some notes on economy of derivation and representation. *MIT Working Papers in Linguistics* 10. 43-74.
- Cinque, G. (1985), Bare quantifiers, quantified NPs and the notion of operator at S-structure. *Rivista di Grammatica Generativa* 11.
- Cinque, G. (1984), A'-bound pro vs. variable. Ms. Università di Venezia.
- Cinque, G. (1989), 'Long' WH-movement and referentiality. Ms. Università di Venezia.
- Clements, G.N., J. McCloskey, J. Maling and A. Zaenen (1983), String-vacuous rule application. *Linguistic Inquiry* 14. 1-17.

- Chung, S. and J. McCloskey (1983), On the interpretation of certain island facts in GPSG. *Linguistic Inquiry* 14. 704-713.
- Dobrovie-Sorin, C. (1988), Clitic doubling, WH-movement and quantification in Romanian. To appear in *Linguistic Inquiry*.
- Engdahl, E. and E. Ejerhed (eds.) (1982), *Readings on Unbounded Dependencies in the Scandinavian Languages*. Umeå: Almqvist & Wiksell.
- Evers, A. (1975), *The Transformational Cycle in Dutch and German*. Doctoral dissertation, Rijksuniversiteit Utrecht. Distributed by Indiana University Linguistics Club.
- Fanselow, G. (1987), *Konfigurationsfähigkeit*. Tübingen: Narr.
- Fanselow, G. (1988), Aufspaltung von NPn und das Problem der freien Wortstellung. *Linguistische Berichte* 114. 91-113
- Fanselow, G. (1989), Scrambling as NP-Movement. in: Grewendorf, G. and W. Sternefeld (eds.).
- Fodor, J.D. (1978), Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9. 427-473.
- Frazier, L. (1987), Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory* 5. 519-559.
- George, L. (1980), *Analogical Generalization of Natural Language Syntax*. Doctoral dissertation, MIT. Cambridge, Massachusetts.
- Grewendorf, G. (1988), *Aspekte der deutschen Syntax*. Tübingen: Narr.
- Grewendorf, G. (1986), *Ergativität im Deutschen*. Ms. Universität Frankfurt. To be published in English by Foris.
- Grewendorf, G. and W. Sternefeld (eds.), *Scrambling and Barriers*. Amsterdam: John Benjamins.
- Haider, H. (1983), Connectedness effects in German. *Groninger Arbeiten zur Germanistischen Linguistik* 23. 82-119.
- Haider, H. (1985), The case of German. in: Toman, J. (ed.), *Studies in German Grammar*. Dordrecht: Foris.
- Haider, H. (1988), Matching projections. in: Cardinaletti, A., G. Cinque and G. Giusti (eds.), *Constituent Structure*. Dordrecht: Foris.
- Hankamer, J. (1973), Unacceptable ambiguity. *Linguistic Inquiry* 4. 17-68.
- Huang, J. (1982a), *Logical Relation in Chinese and the Theory of Grammar*. Doctoral dissertation, MIT Cambridge, Mass.
- Jackendoff, R. and P. Culicover (1971), A reconsideration of dative movements. *Foundations of Language* 7. 397-412.
- Kayne, R. (1983), Connectedness. *Linguistic Inquiry* 14 223-249.
- Kayne, R. (1989), Null subjects and clitic climbing. in: Jaeggli, O. and K. Safir (eds.), *The Null Subject Parameter*. Dordrecht: Reidel.
- Koopman, H. (1983), ECP effects in main clauses. *Linguistic Inquiry* 14 346-350.

- van de Koot, H. (1988), The vacuous movement hypothesis, superiority and the ECP. in: de Haan, G. and W. Zonneveld (eds.), *Formal Parameters of Generative Grammar*. IV Yearbook 1988. Dordrecht.
- Koster, J. (1978), *Locality Principles in Syntax*. Dordrecht: Foris.
- Koster, J. (1986), *Domains and Dynasties: The Radical Autonomy of Syntax*. Dordrecht: Foris.
- Kratzer, A. (1989), Stage-level and individual-level predicates. Ms. University of Massachusetts, Amherst.
- Lasnik, H. and M. Saito (1984), On the nature of proper government. *Linguistic Inquiry* 15. 235-289.
- Lebeaux, D. (1989), Parameter-setting, the acquisition sequence, and the form of grammar: The composition of phrase structure. Talk presented at the GLOW-conference 1989, Utrecht.
- Lötscher, A. (1978), Zur Verbstellung im Zürichdeutschen und in anderen Varianten des Deutschen. *Zeitschrift für Dialektologie und Linguistik* 45. 1-29.
- McElree, B. and T. Bever (1989), The psychological reality of linguistically defined gaps. *Journal of Psycholinguistic Research* 18. 21-35.
- Nicol, J. (1988), Re-activating antecedents of empty categories during parsing. Ms. Department of Psychology, University of Arizona.
- Nicol, J. and D. Swinney (1989), The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research* 18. 5-19.
- Pesetsky, D. (1986), WH-in-situ: Movement and unselective binding. in: Reuland, E. and A. ter Meulen (eds.), *The Representation of (In)definiteness*. Cambridge, Massachusetts: MIT Press.
- Pollock, J.-Y. (1989), Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20. 365-424.
- Reis, M. (1985), Satzeinleitende Strukturen im Deutschen. über COMP, Haupt- und Nebensätze, W-Bewegung und die Doppelkopfanalyse. in: Abraham, W. (ed.), *Erklärende Syntax des Deutschen*. Tübingen: Narr.
- Reis, M. and I. Rosengren (1988), WH-imperatives?! *Sprache und Pragmatik* 10, Lund. 1-49.
- van Riemsdijk, H. (1989), Movement and regeneration. in: Beninca, P. (ed.), *Dialect Variation and the Theory of Grammar*. Dordrecht: Foris.
- Rizzi, L. (1982), *Issues in Italian Syntax*. Dordrecht: Foris.
- Rizzi, L. (1986), Null-objects in Italian and the theory of pro. *Linguistic Inquiry* 17. 501-557.
- Rizzi, L. (1987), Relativized minimality. Ms. Universite de Geneve.
- Rizzi, L. (1989), *Relativized Minimality*. Ms. Universite de Geneve. To be published by MIT-Press.
- von Stechow, A. und W. Sternefeld (1988), *Bausteine syntaktischen Wissens. Ein Lehrbuch der generativen Grammatik*. Opladen: Westdeutscher Verlag.

- Stowell, T. (1981), *Origins of Phrase Structure*. Doctoral dissertation, MIT. Cambridge, Massachusetts.
- Tappe, T. (1984), On infinitival clauses without COMP. in: de Geest, W. and Y. Putseys (eds.), *Sentential Complementation*. Dordrecht: Foris.
- Taraldsen, T. (1980), On the Nominative Island Condition, vacuous application and the *that*-trace filter. *Indiana University Linguistics Club*.
- Thiersch, C. (1978), *Topics in German Syntax*. Doctoral dissertation, MIT. Cambridge, Massachusetts.
- Travis, L. (1984), *Parameters and Effects of Word Order Variation*. Doctoral dissertation, MIT. Cambridge, Massachusetts.
- Webelhuth, G. (1988), A universal theory of scrambling. in: Rosen, V. (ed.), *Papers from the Tenth Scandinavian Conference of Linguistics*, vol. 2. Bergen: Department of Linguistics and Phonetics.
- Williams, E. (1978), Across-to-board rule application. *Linguistic Inquiry* 9. 31-43.