Imperfect Checking[1]

Gereon Müller
Universität Stuttgart

Abstract

This article presents a unified solution to two long-standing problems concerning successive cyclicity. First, under the minimalist program of Chomsky (1995), a trigger for successive-cyclic *wh*-movement to [–wh] SpecC positions must be identified. Second, as noted by den Besten & Webelhuth (1990), there is a strict prohibition against unbound intermediate traces that result from a combination of successive-cyclic *wh*-movement and remnant topicalization in German. The key proposal of this article is a new concept of "imperfect checking" of [+wh] and [–wh] features, which triggers successive-cyclic *wh*-movement. Moreover, as a parametrized reflex of successive cyclicity, extraposition can be forced. Obligatory extraposition accounts for the prohibition against unbound intermediate traces (via the strict cycle, economy, and general properties of features). In addition, the theory of imperfect checking yields interesting consequences for partial *wh*-movement and infinitives in German.

1. Introduction

The main goal of this paper is to provide a unified solution to two seemingly unrelated problems that arise with successive-cyclic movement. The analysis is carried out on the basis of the minimalist program developed in Chomsky (1995). The first problem is theory-internal: Assuming that each movement operation must be triggered by a feature, what triggers successive-cyclic *wh*-movement to [–wh] SpecC positions? The second problem, in contrast, is an empirical one: Given that remnant topicalization in German may create unbound traces, why is it that these unbound traces must always be initial ones, i.e., why can remnant movement not create unbound intermediate traces that result from successive-cyclic movement?

To answer the first question, I will argue that the theory of feature checking is not restricted to the concepts of perfect feature checking and feature mismatch (which cancels the derivation), as is standardly assumed, but, in addition, makes available a concept of what I would like to call "imperfect checking." Imperfect checking differs from perfect checking in requiring identity of features, but not identity of feature values. Imperfect checking is what takes place in cases of *wh*-movement to $\text{SpecC}_{[-wh]}$ (where a [+wh] feature of a *wh*-phrase is checked with a [–wh] feature of C). This will be shown to induce a feature strengthening operation. Feature strengthening in turn triggers (a) the next step of successive-cyclic *wh*-movement,

[1]For helpful discussion, I would like to thank Josef Bayer, Daniel Büring, Jonathan Bobaljik, Gisbert Fanselow, Danny Fox, Wolfgang Sternefeld, Anita Steube, an anonymous reviewer for *The Linguistic Review*, and audiences at the 21. GLOW colloquium, University of Tilburg (April 1998), and at Leipzig University (May 1998).

and (b) obligatory CP extraposition in languages like German and Basque, which can thus be viewed as a reflex of successive cyclicity. This latter operation will then be shown to be responsible for the prohibition against unbound intermediate traces created by remnant movement, whereby the second question posed above is answered.

I will proceed as follows. Section 2 briefly introduces remnant movement constructions in German from a minimalist perspective and highlights the problem posed by unbound intermediate traces. Section 3 turns to successive cyclicity and motivates an approach in terms of imperfect checking and feature strengthening. Section 4, then, puts the two things together and shows how the new approach to successive cyclicity directly accounts for the prohibition against unbound intermediate traces, without recourse to construction-specific assumptions. Section 5 shifts the focus from finite CPs to infinitives, which behave differently with respect to both unbound traces and successive cyclicity. Finally, section 6 draws a conclusion and points out some further consequences of the approach developed here.

2. Remnant Movement

2.1. Remnant Movement in the Principles and Parameters Approach

The idea that German topicalization examples like those in (1) should be analyzed as instances of remnant movement goes back to Thiersch (1985) and den Besten & Webelhuth (1987).

(1) a. [$_{VP_2}$ t_1 Gelesen] hat das Buch$_1$ keiner t_2
read has the book$_{acc}$ no-one$_{nom}$
'No-one read the book.'

b. [$_{\alpha_2}$ t_1 Zu lesen] hat das Buch$_1$ keiner t_2 versucht
to read has the book$_{acc}$ no-one$_{nom}$ tried
'No-one tried to read the book.'

On this view, VP_2 topicalization in (1-a) and topicalization of the coherent infinitive α_2 in (1-b) is preceded by a scrambling operation that removes NP_1 from XP_2, thereby turning XP_2 into a remnant category. Consequently, remnant topicalization creates an unbound trace t_1. The remnant movement approach to "incomplete category fronting" as in (1) has been adopted and further substantiated in Stechow & Sternefeld (1988, ch. 12), den Besten & Webelhuth (1990), Huang (1993), Bayer (1993), Grewendorf (1994), Grewendorf & Sabel (1994), Sabel (1995), and Müller (1998), and in what follows, I will presuppose that it is essentially correct. However, despite the substantial amount of evidence in support of remnant movement, certain aspects of the concept have been inherently problematic under representational approaches to movement phenomena; see, e.g., Haider (1991, 28). Most notably, if one looks only at their surface structures, the examples in (1) are not straightforwardly distinguishable from ill-formed examples involving syntactic lowering, as in (2-a), or freezing effects, as in (2-b).

(2) a. *John asked t_1 [$_{CP}$ who$_1$ Mary saw Bill]
b. *Who$_1$ did you say that [$_{NP_2}$ pictures of t_1] were stolen t_2 ?

In a representational approach, (2-a) is standardly excluded by the Proper Binding Condition: Overt *wh*-lowering is illicit because it produces an unbound trace (cf. Fiengo 1977 and May 1977, among many others). The freezing effect in (2-b) is explained in Browning (1991) by invoking a theory of movement according to which the locality requirements for overt movement traces are checked by looking at S-structure (cf. Lasnik & Saito 1984, Chomsky 1986): At S-structure, the derived subject NP_2 occupies a position that is not L-marked, and hence it qualifies as a barrier for t_1 – the fact that NP_2 is transparent for extraction in its VP-internal base position does not matter. Clearly, by the same reasoning we should then expect the remnant movement examples in (1) to be excluded as well: Not only does an unbound trace occur here; what is more, that trace is also included in an S-structure barrier (the topicalized XP_2s in (1) are not in L-marked positions anymore). Thus, to reconcile the very existence of remnant movement with an account of lowering and freezing effects, otherwise unmotivated assumptions seem to be called for (cf. den Besten & Webelhuth 1990, Rizzi 1990, and Grewendorf & Sabel 1994).

2.2. Remnant Movement in the Minimalist Program

The theory of movement adopted in the minimalist program is inherently derivational. In particular, the account of illicit lowering and freezing effects does not rely on the concept of S-structure anymore. Interestingly, it turns out that these two problems that remnant movement constructions pose for a representational approach now cease to exist. Crucially, the ill-formed constructions in (2) involve illegitimate movement steps in the derivation, whereas the well-formed constructions in (1) involve only legitimate movement steps: The fact that the resulting representations in (1) look like illicit lowering and freezing configurations on the surface is irrelevant.[2] Thus, in a strictly derivational approach to movement, (2-a) can be excluded by the requirement that all movement is raising, and there is little reason to assume that something like the Proper Binding Condition exists at all. Clearly, all movement is raising in (1), and the availability of this kind of construction is not at all surprising.

Furthermore, (2-b) can be accounted for by the interaction of the Barriers Condition (BC) and the Strict Cycle Condition (SCC); on this view, the concept of S-structure can (and should) be dispensed with. For concreteness, let us assume that there is a derivational constraint that precludes movement across a barrier (cf. Chomsky 1986; 1995, 264 & 328; 1998, 13).[3]

---

[2]Cf. Chomsky (1998, 15-16), who remarks that, after a movement operation has taken place, "the raised element c-commands its trace in the original position ... Further operations might lead to violation of the relation, as in ... independent XP-dislocation [i.e., remnant XP movement, G.M.], in which the step-by-step locality and c-command relations ... are obliterated."

[3]Attempts have been made to derive BC from more basic concepts; cf. Kitahara (1994) and Takahashi (1994), among others. However, for present purposes it may suffice to adopt BC as

(3) *Barriers Condition (BC):*
Movement must not cross a barrier.

Assuming with Cinque (1990) that all XPs in derived positions are non-L-marked, this means that XP movement invariably turns an XP into a barrier, even if it is transparent for extraction in situ.[4] Suppose next that, in addition to BC, there is a Strict Cycle Condition (SCC) that takes the following form.[5]

(4) *Strict Cycle Condition (SCC):*
A strong feature F cannot be checked in a domain α if α is embedded in a domain β, and a strong feature has been already been checked in β.

The interaction of BC and SCC straightforwardly captures freezing effects as in (2-b) (cf. Browning 1989 and Collins 1994): If NP_2 raising precedes *wh*-movement of NP_1, SCC is respected (the second movement operation ends up in a higher position than the first one), but BC is violated (extraction from NP_2 takes place at a stage in the derivation at which NP_2 is in subject position, which is non-L-marked). If, on the other hand, *wh*-movement of NP_1 precedes NP_2 raising, BC is respected (NP_2 is still in object position, hence L-marked, when extraction takes place), but SCC is violated (the EPP feature cannot be checked in the embedded TP domain because a strong *wh*-feature has already been checked in the higher CP domain). Thus, there is no well-formed derivation left, and the freezing effect is explained.[6] In contrast, no such freezing effect shows up in remnant movement constructions like those in (1): If, e.g., NP_1 scrambling in (1-a) precedes VP_2 topicalization, both BC and SCC are respected: All movement takes place from XPs in L-marked (i.e., complement) positions, and VP_2 topicalization ends up in a higher position than the preceding

---

such, and to leave open the question of whether a genuine derivation of its effects is possible.

[4]Note that this does not imply that only XPs in derived positions can act as barriers. In fact, it will later (in subsection 4.2) become relevant that even an L-marked CP is a barrier for movement that does not start from CP's specifier. See Baker (1988), Sportiche (1989) and Müller & Sternefeld (1993; 1995) for notions of barrier that accomplish this in a relatively homogeneous way.

[5]This formulation is close to the one given in Chomsky (1995, 233-234), but differs in subtle respects. These are related to the fact that the movement theory to be developed below presupposes that strong features can occur and be checked in embedded, non-root configurations, as long as no strong feature in a higher position has been checked so far. Note also that the general ban on lowering follows from SCC, as far as overt movement is concerned, if we assume that all pre-Spell-Out movement is triggered by strong features; see below.

[6]Browning (1991) observes that there is a third type of derivation that is a priori possible and must be excluded, one in which *wh*-extraction takes place first (thereby fulfilling BC) but ends up in a VP-adjoined position, followed by NP raising to SpecT and, finally, a second instance of *wh*-movement from the VP-adjoined position to SpecC. However, such a derivation can readily be blocked as an instance of improper movement by invoking either a uniformity condition on chain formation, or (as first suggested by Collins 1994) economy considerations. The two constraints on improper movement envisaged in (33) (uniformity) and (34) (economy) in subsection 4.2 below would both suffice to exclude the illicit derivation at hand.

NP_1 scrambling operation.[7]

To sum up, it seems that the strictly derivational theory of movement that underlies the minimalist program avoids a number of problems for the remnant movement approach that arose under more representational theories – remnant movement structures do not "look" right, but viewed from a derivational perspective, there is nothing wrong with them (and this change of perspective may also partly be responsible for the rich use of remnant movement structures that is made in Kayne (1998) and related work).[8] Still, there are certain peculiar restrictions on remnant movement constructions that minimalist movement theory as such does not seem to say anything about, and one of these is the strict prohibition against unbound intermediate traces in German remnant movement constructions.

2.3. The Ban on Unbound Intermediate Traces

2.3.1. The Phenomenon

Note first that complete (i.e., non-remnant) finite CPs and VPs with an internal V-CP structure can be topicalized in German:

(5) a. [$_{CP_2}$ Daß Fritz Claudia liebt] hat er gesagt
that Fritz$_{nom}$ Claudia$_{acc}$ loves has he said
'He said that Fritz loves Claudia.'

b. [$_{VP_3}$ Gesagt [$_{CP_2}$ daß er sie liebt]] hat er gestern t_3
said that he$_{nom}$ her$_{acc}$ loves has he yesterday
'Yesterday, he said that he loves her.'

Second, as noted by Fanselow (1987), Bayer (1990), and Müller & Sternefeld (1993), topicalization of complement XPs across a *wh*-island normally induces only weak illformedness (a subjacency-like effect) in German.[9] Non-remnant finite argument CPs and complex VPs with a V-CP structure behave as expected, given their complement status:

---

[7]Of course, the reverse application of movement operations results in a derivation that violates both BC and SCC. But since there is one derivation that is well formed, and since economy considerations do not force us to prefer one derivation over the other, this is unproblematic.

[8]Still, while concerned with constructions of a very different type, Chomsky (1995, 365) introduces a condition that would also block all instances of remnant movement. This condition (200) would therefore have to be either given up, or revised accordingly. In fact, a modification can be shown to be required on independent grounds; cf. Müller (1998, ch. 1).

[9]I will not have anything new to say about the nature of the constraint that is responsible for weak ungrammaticality with topicalization from *wh*-islands, or why it is that *wh*-extraction from *wh*-islands always results in strict ungrammaticality, even with arguments; but cf. footnote 38.

(6) a. ?[$_{CP_2}$ Daß Fritz Claudia liebt] weiß ich nicht [$_{CP}$ ob er
that Fritz$_{nom}$ Claudia$_{acc}$ loves know I not whether he
gesagt hat]
said has
'I do not know whether he said that Fritz loves Claudia.'

b. ?[$_{VP_3}$ Gesagt [$_{CP_2}$ daß er sie liebt]] weiß ich nicht [$_{CP}$ ob
said that he$_{nom}$ her$_{acc}$ loves know I not whether
er t_3 hat]
he has
'I do not know whether he said that he loves her.'

Third, remnant VPs and remnant infinitives as they show up in (1) can in principle be topicalized across a *wh*-island that is erected by the antecedent of the unbound trace (i.e., in these cases, the remnant XPs contain an unbound *wh*-trace, rather than an unbound scrambling trace, as in (1)). The resulting sentences exhibit only a weak subjacency-like effect, just like the sentences in (6).

(7) a. ?[$_{VP_2}$ t_1 Gelesen] weiß ich nicht [$_{CP}$ was$_1$ er t_2 hat]
read know I not what he has
'I do not know what he read.'

b. ?[$_{\alpha_2}$ t_1 Zu lesen] weiß ich nicht [$_{CP}$ was$_1$ er t_2 versucht hat]
to read know I not what$_{acc}$ he tried has
'I do not know what he tried to read.'

Interestingly, even though complete finite CPs and VPs with a V-CP structure can be topicalized from a *wh*-island, and even though remnant movement is in principle possible from a *wh*-island erected by the antecedent of the unbound trace, it turns out that remnant finite CPs and remnant VPs with a V-CP structure that contain an unbound *intermediate* trace resulting from successive-cyclic *wh*-movement (which in turn is required for locality reasons) cannot be topicalized across a *wh*-island created by the antecedent of the unbound trace. As shown in (8), the result is much worse than we would expect if only a subjacency-like effect were involved (as a matter of fact, sentences of this type can hardly be made sense of by unknowing speakers).

(8) a. *[$_{CP_2}$ t'_1 Daß Fritz t_1 liebt] weiß ich nicht [$_{CP}$ wen$_1$ er gesagt hat]
that Fritz$_{nom}$ loves know I not who$_{acc}$ he said has
'I do not know who he said that Fritz loves.'

b. *[$_{VP_3}$ Gesagt [$_{CP_2}$ t'_1 daß Fritz t_1 liebt]] weiß ich nicht [$_{CP}$ wen$_1$ er
said that Fritz$_{nom}$ loves know I not who$_{acc}$ he
t_3 hat]
has
'I do not know who he said that Fritz loves.'

Thus, the following generalization seems to be descriptively adequate (cf. den Besten & Webelhuth 1990, Fanselow 1993, Grewendorf 1994, Bayer 1996, and Müller 1998):

(9) *Generalization*:
Remnant movement in German may not create unbound intermediate traces.

The question arises of how this generalization can be derived. The first thing to note is that it does not follow from (the interaction of) BC and SCC: Nothing so far (modulo the weak *wh*-island effect) rules out a derivation of (8-a) and (8-b) that proceeds by two instances of *wh*-movement across VP_3 and CP_2 (while these are still in situ, in an L-marked position), followed by topicalization of either CP_2 (cf. (8-a)) or VP_3 including CP_2 (cf. (8-b)) to a higher position (in accordance with SCC).

Previous approaches to the phenomenon include den Besten & Webelhuth (1990), who introduce a condition which states that traces in non-argument positions cannot be reconstructed, and Grewendorf (1994), who suggests a constraint according to which non-trivial chains cannot be reconstructed (both proposals rely on the idea that the Proper Binding Condition holds at LF, and that unbound traces created by remnant movement must be bound by that level). However, certain empirical problems notwithstanding (cf. Müller 1998, ch. 6), it seems clear that these approaches, in their explicit recognition of a different behaviour of bound and unbound traces, and in their implicit acknowledgement of a different behaviour of initial and intermediate unbound traces, depend on construction-specific assumptions of a type that one would ideally want to do without – in fact, for the most part these approaches can arguably be viewed as elegant reformulations of (9).

Still, before I set out to develop a less surface-oriented approach, it should be noted that den Besten & Webelhuth (1990) also discuss the ill-formed constructions in (10-ab) that involve postposition stranding and subextraction (*was für* split), respectively. They propose that these constructions can be handled by the same constraint against unbound traces in non-argument positions.

(10) a. *[$_{PP}$ t_1 Mit]$_2$ hat da$_1$ keiner t_2 gerechnet
with has there no-one counted
'No-one could know that.'

b. *[$_{NP}$ t_1 [$_{PP}$ Für Bücher]]$_2$ weiß ich nicht [$_{CP}$ was$_1$ er t_2 gelesen hat]
for books know I not what he read has
'I do not know what kinds of books he read.'

On their view, t_1 in (10-ab) is a trace in a non-argument position. The approach to be developed in what follows will have nothing to say about the illformedness of these examples. However, closer inspection reveals that matters are far more complex. On the one hand, it turns out that the constructions in (10-ab) cannot be ruled out in the same way as those in (8) even under den Besten & Webelhuth's own assumptions; e.g., nothing in their analysis excludes topicalization of the trace-less PP in (10-b) (instead of the whole NP), in which case there would be no offending unbound trace present. And on the other hand, alternative explanations of the illformedness of (10-ab) that are not related to remnant movement are readily available. Thus, as concerns (10-a), there is a general ban on bare prepositions in SpecC in German that is independently motivated by the impossibility to front a separable prefix in verb-particle constructions in German (assuming with much recent literature that

the particle forms an intransitive PP in the base):[10]

(11) *[$_{PP}$ Mit]$_2$ hat gestern keiner t_2 gespielt
with has yesterday no-one played
'No-one joined in the game yesterday.'

Furthermore, closer inspection reveals that (10-b) simply illustrates a subcase of a peculiar restriction on subextraction in German; see Müller (1998, 15-17, 267-269, & 316-320) for discussion of both issues.

That said, I will now address the question of how the ban on unbound intermediate traces is to be accounted for.[11]

2.3.2. Towards a Solution

In what follows, I will pursue the hypothesis that the prohibition against unbound intermediate traces is related to the position occupied by CP_2 in (8). Two possibilities arise: Either V-CP order is base-generated in SOV languages like German (cf. Koster (1987), Webelhuth (1992), Zwart (1993), Haider (1994), and Bayer (1996), to name but a few), or V-CP order is transformationally derived from a base-generated CP-V order by CP extraposition. The first view can probably be considered to be the predominant one nowadays. However, Büring & Hartmann (1997) have presented a number of arguments against it, and for a uniform CP-V base order in German, with subsequent rightward movement of CP. I will adopt this latter view

---

[10]Topicalization of a bare preposition in verb-particle constructions yields somewhat better results if the preposition is semantically independent, i.e., if the denotation of the verb-particle construction can be determined fully compositionally; cf. (i-a). Interestingly, the same effect can be seen with remnant PP topicalization; cf. (i-b) vs. (i-c).

(i) a. ?[$_{PP}$ Zu]$_2$ hat sie die Tür t_2 gemacht
shut has she the door made
'She closed the door.'
b. *[$_{PP}$ t_1 Für]$_2$ hat da_1 keiner etwas t_2 gekonnt
for has there no-one something could
'Nobody was responsible.'
c. ??[$_{PP}$ t_1 Für]$_2$ hat da_1 keiner t_2 gestimmt (gegen schon)
for has there no-one voted (against indeed)
'No-one voted for it (but someone voted against it).'

The intermediate status of examples like (i-c) can be taken as a further argument against assimilating the less rigid ban on remnant PP topicalization and the strict prohibition against unbound intermediate traces created by successive-cyclic movement.

[11]A very simple explanation is given by Fanselow (1993), who argues that the examples in (1) should be reanalyzed as involving topicalization of the bare verb or infinitive, with no unbound trace present, and that the illformedness of the examples in (8) in turn shows that unbound traces are never permitted in German in the first place (since there is no way to reanalyze these constructions as not involving *wh*-traces). This approach is compelling in its simplicity; however, it is in conflict with the substantial amount of evidence that has been accumulated in support of remnant movement structures in German; cf. the references in subsection 2.1.

in what follows. Indeed, I would like to contend that, under the view that V-CP structures are base-generated, it is impossible for principled reasons to account for the contrast between ill-formed examples like (8-a) and (8-b) on the one hand, and (halfway) well-formed examples like (7-a) and (7-b) on the other, without invoking construction-specific assumptions that more or less explicitly differentiate between bound and unbound traces, as well as between intermediate and initial traces. In contrast, I will argue that if one assumes CP-V order to be base-generated, and V-CP order to be derived by extraposition, a non-construction-specific account of the prohibition against unbound intermediate traces can be given.

As a matter of fact, it turns out that one type of example covered by the generalization in (9) follows directly from the interaction of BC and SCC, as a freezing effect, once we adopt the idea that V-CP order is transformationally derived. To see this, consider again (8-b), now enriched by a pre-verbal CP trace:

(12) *[$_{VP_3}$ t_2 Gesagt [$_{CP_2}$ t'_1 daß Fritz t_1 liebt]] weiß ich nicht [$_{CP}$ wen_1 er t_3
said that Fritz$_{nom}$ loves know I not who$_{acc}$ he
hat]
has
'I do not know who he said that Fritz loves.'

There are four movement operations to be considered, and it remains to be shown that no order of application meets all movement constraints. It is clear from the start that both *wh*-movement operations must precede CP extraposition; otherwise, BC will be violated (CP occupies a non-L-marked position after extraposition). Similarly, both *wh*-movement operations must precede VP topicalization; *wh*-movement from a topicalized VP will violate both BC (VP occupies a non-L-marked position after topicalization) and SCC (the landing site of *wh*-movement is lower than the landing site of the preceding topicalization operation). Thus, the only two remaining derivations start with two applications of *wh*-movement, followed by CP extraposition and VP topicalization, in either order. However, both these derivations invariably violate SCC. If CP extraposition applies third, before VP topicalization, then SCC is violated in the third step (the VP_3-adjoined landing site of CP_2 is dominated by the C domain in which the [+wh] feature wen_1 has been checked). If CP extraposition applies last, after VP topicalization, then SCC is violated in the last step (the VP_3-adjoined landing site of CP_2 is dominated by the root domain in which the topic feature – or whatever feature triggers topicalization – of VP_3 itself has been checked). Hence, there is no legitimate derivation left for an example like (12), and it is correctly ruled out without resort to a specific constraint on unbound traces. Essentially, on this view the problem with (12) is that the landing site of CP extraposition is too low to avoid a freezing effect; thus, (12) receives roughly the same analysis as the typical freezing example in (2-b).[12]

[12]The question arises of how this freezing effect is avoided in ordinary cases of extraction from finite CPs in German. I will address this issue below; but for the time being, it is clear that CP extraposition will have to target a higher position than a VP-adjunction site in that case, in

Thus, of the two types of examples covered by (9), one is now accounted for by assuming that V-CP structures are derived by CP extraposition. However, at this point it is still unclear what rules out bare CP topicalization as in (8-a). Furthermore, the CP extraposition hypothesis may have contributed to the solution of one problem, but it introduces a new problem: If CP extraposition to VP is fatal in (12), we expect that the structure can be saved by simply leaving CP in its pre-verbal in situ position. As shown in (13), this prediction is not borne out; (13) is just as impossible as (12) is, even though there can be no freezing effect at work here.[13]

(13) *[$_{VP_3}$ [$_{CP_2}$ t'_1 Daß Fritz t_1 liebt] gesagt] weiß ich nicht [$_{CP}$ wen$_1$ er t_3
that Fritz$_{nom}$ loves said know I not who$_{acc}$ he
hat]
has
'I do not know who he said that Fritz loves.'

To solve these two remaining problems and thereby derive (9) in toto, I will take a closer look at the mechanics of successive-cyclic movement in German in the next section, and I will identify obligatory extraposition as a reflex of successive cyclicity in German that is ultimately responsible for the illformedness of (8-a) and (13).

3. Successive Cyclicity

3.1. Successive Cyclicity in the Minimalist Program

Given Last Resort (LR) in (14) (cf. Chomsky 1995, 296-297), syntactic movement must be triggered by feature checking (in addition, given the Mininimal Link Condition (MLC) which is integrated into LR as requirement (b), feature checking must be highly local).

(14) *Last Resort (LR):*
α can raise to position β only if (a) and (b) hold:
a. α checks a feature F in β.
b. There is no γ with an unchecked feature F that is closer to β (MLC).

Under this assumption, successive-cyclic *wh*-movement as in (15-a) poses a problem. For reasons of locality, the two intermediate adjunct traces are necessary in (15-a), but it is not immediately clear whether the [–wh] SpecC positions that these traces occupy (cf. (15-b)) are positions in which the *wh*-phrase checks a feature on the way to its [+wh] SpecC target position in the root clause.

(15) a. How$_1$ do you think [$_{CP}$ t''_1 (that) John said [$_{CP}$ t'_1 (that) Bill fixed the

accordance with SCC, an option that is not available in (12) since VP topicalization signals that CP has not left the VP here.

[13]Note that the markedness that is usually associated with pre-verbal finite CPs in German (see below) cannot solely be held responsible for the sharp ungrammaticality evident in (13).

car t_1]] ?

b. ... $C_{[+wh]}$ ... $C_{[-wh]}$ ... $C_{[-wh]}$ ... $XP_{[+wh]}$

In view of this, two different strategies have been pursued. First, Chomsky (1993) proposes that movement is to be analyzed as a somewhat more complex operation "Form Chain." On this approach, $XP_{[+wh]}$ raises to $C_{[+wh]}$ in one step in (15-b), and the two intermediate traces are subsequently inserted in the specifiers of the intervening $C_{[-wh]}$ nodes. Second, it has been suggested that feature checking does indeed take place in the domains of [–wh] C nodes in the course of *wh*-movement in (15-b), as required by LR (cf. Ferguson & Groat 1994, Fanselow & Mahajan 1996, Sabel 1996, and Collins 1997, among others). According to these analyses, there is a feature F that is distinct from the [+wh] feature that the *wh*-phrase checks in its [+wh] target position; this feature F is checked with intervening $C_{[-wh]}$ nodes. There is some evidence that the second type of approach is indeed more on the right track. For one thing, it is preferable on conceptual grounds in a derivational grammar (the concept of Form Chain is essentially a representational one and, furthermore, inherently acyclic since movement followed by trace insertion involves referring back to an embedded structure). For another, intermediate feature checking seems to be more in accord with the existence of reflexes of successive cyclicity in a number of languages. Thus, the choice of complementizer (*aL* vs. *goN*) in Modern Irish (cf. McCloskey 1979) that depends on whether or not a C domain has been affected by movement lends itself most naturally to an account in terms of intermediate feature checking. The same goes for the occurrence of obligatory V-to-C raising with (certain types of) *wh*-phrases in Spanish (cf. Torrego 1984, Baković 1998) and Basque (cf. Ortiz de Urbina 1989); for the selection of subject pronouns by C in Ewe (cf. Collins 1994); for the presence vs. absence of tonal downstep in Kikuyu (cf. Clements et al. 1983); for *meN* deletion in Bahasa Melayu (colloquial Singapore Malay) (cf. Cole & Hermon 1998); and so on.

However, this approach also faces a number of problems. First, the nature of the "non-[+wh]" feature F in question remains notoriously unclear in the literature.[14] Second, and perhaps more importantly, it is clear that F must be optionally present on a [–wh] C – a derivation will be cancelled if F is present on C, and *wh*-movement

[14]Suggestions include "[wh] of some sort" (vs. "real" [wh]), [+wh] (vs. "real" [Q]), [Op] (vs. "real" [+wh]), etc., all of which might be taken to mean that, if we assume that the feature that the *wh*-phrase checks in its ultimate target position is [+wh], the feature that is checked in embedded [–wh] SpecC positions is a "pseudo"-[+wh] feature. This, in turn, is a notion that hardly meets the minimalist requirement of being basic or "primitive;" rather, it looks like an artefact of the theory. Another suggestion that can be found (cf. Fanselow & Mahajan 1996) is that the "non-[+wh]" feature F on [–wh] C nodes is in fact a [D] feature, which is in line with the hypothesis that, ideally, only categorial features should be able to attract. But, if applied naively, this approach is far too strict, since it predicts that only nominal categories can undergo successive-cyclic movement, and that, assuming the MLC, only subjects should be able to be affected by *wh*-movement in the first place. Consequently, additional assumptions are then called for, which in turn seem to be motivated mainly by theory-internal considerations (e.g., that *wh*-moved PPs and APs can be viewed as nominal categories bearing a [D] feature, or that the MLC does not hold).

to C does not apply (cf. Collins 1997, 106). Thus, F must be instantiated on exactly n [–wh] C nodes in the numeration if a *wh*-phrase is to enter n F checking relations with [–wh] C nodes in the derivation, on its way to the [+wh] target position. But this latter information is not yet available in the numeration. Hence, this approach faces a severe problem of overgeneration.[15] I would like to conclude from this that while there is some evidence for checking of a feature F in constructions involving successive-cyclic movement, the actual proposals that have been made are problematic because of (a) the unclear nature and (b) the optional presence of F.

3.2. Feature Strengthening

To solve the first problem (a), I would like to suggest that there is nothing mysterious about the "non-[+wh] feature F," and that it is superfluous to enrich by inventory of basic features in order to cope with successive cyclicity – what triggers movement to [–wh] SpecC positions is the feature [–wh] itself, which has to show up on declarative C nodes for independent selectional reasons. This assumption will be shown to also cover problem (b): [–wh] is not an optional feature on declarative C. It is obligatorily present, but it may be weak or strong. On this view, once started, successive-cyclic movement to SpecC_n is triggered by a feature strengthening process that arises as a result of imperfect checking of [+wh] and [–wh] features in SpecC_{n-1}, and that also induces CP extraposition in German.

More specifically, I would like to contend that the feature [±wh] plays a dual role on C.[16] On the one hand, it must be checked against a corresponding feature of the c-commanding matrix predicate, and on the other hand, it can be checked against a corresponding feature of a c-commanded *wh*-phrase; i.e., [±wh] on C participates both in selection and in *wh*-movement.

Let me begin with selection. With D-structure gone, it strikes me as plausible to encode the selection relation between a matrix predicate and the CP it embeds

---

[15] One might think that this problem arises throughout, and is inherent to the idea of a numeration – the vast majority of a priori possible numerations would, on this view, not end up in well-formed derivations. However, such an argumentation overlooks the fact that in many cases, extremely simple strategies are readily available that help reducing the number of numerations that are doomed to result in ill-formed derivations from the very beginning. For instance, it is clear that if three transitive verbs are chosen in one numeration, then there must be six arguments present as well in order to prevent a subsequent violation of the θ-criterion. Similarly (abstracting away for the time being from multiple questions; see below), if there are, say, two *wh*-phrases in the numeration, then there should also be two $\text{C}_{[+wh]}$s present; for every T, there should be one V, etc. There is no reason why such information, which is locally available and does not presuppose look-ahead knowledge about what is going to happen in the derivation, should not be assumed to be present during numeration formation, so as to reduce the overgeneration problem. Also see Stabler (1996, ch. 5) for related considerations (cf. in particular the discussion of count invariants in his section 5.4.3). – Crucially, no such strategy is available with the putative F feature responsible for successive-cyclic movement, which sheds doubt on its existence.

[16] For somewhat related considerations with respect to [D] features, albeit with radically different consequences, see Fanselow & Mahajan (1996).

as checking of the selection feature, i.e., [±wh]; the predicate checks the type of its clausal argument, in complete analogy to Case feature checking with nominal arguments. Thus, a predicate like *think* bears a [–wh] feature that must be checked by a matching [–wh] feature of the embedded C by movement; a predicate like *ask* bears a [+wh] feature that must be checked by a [+wh] feature of the embedded C by movement; and a predicate like *know* may bear a [–wh] or a [+wh] feature, which must be checked by a C feature with the matching feature value. Clearly, [±wh] on the matrix predicate is uninterpretable. In contrast, assuming a Karttunen (1977) style semantics of questions, the [+wh] feature of C is interpretable: It introduces a propositional variable that ultimately ensures that questions denote sets of propositions, rather than propositions (cf. Heim 1992, Dayal 1994, Beck 1996, and Stechow 1996, among others.) It seems reasonable to postulate that [–wh] on C, which indicates sentence mood, is interpretable as well; cf. Chomsky (1995, 292). Turning now to strength, I will assume that [±wh] on C is weak, and that, in contrast, [±wh] on the matrix predicate may be strong or weak in German. If [±wh] on the matrix predicate is weak, feature raising from C to the matrix predicate takes place at LF; if [±wh] on the matrix predicate is strong, it overtly attracts [±wh] of the embedded C. C incorporation not being an option in German, overt [±wh] feature raising must pied pipe the whole CP, and I would like to suggest that this is what underlies argument CP extraposition in German: The checking domain of [±wh] on a matrix predicate is confined to right-adjoined positions. The fundamental difference in checking positions for NPs and CPs evident here seems to be directly induced by the data and therefore irreducible; also compare the classic difference between the Case Filter (for NP) and the Case Resistance Principle (for S'/CP) in Stowell (1981).[17] Argument CP extraposition, then, can target a VP-right-adjoined position (cf. (5-b), (6-b)).[18] If we assume that [±wh] may be instantiated not only on a matrix V itself, but on any head of V's extended projection (cf. Grimshaw (1991)), CP extraposition can also target a higher XP-adjoined position. For concreteness, I will assume that, in German, the extended projection of V comprises vP, TP, and TopP, with Top being the highest verbal functional head of a clause, and the landing site of V/2 movement; cf. Müller & Sternefeld (1993) and Zwart (1993). The clause structure presupposed throughout is given in (16):

(16) $[_{\mathrm{CP}}$ – C $[_{\mathrm{TopP}}$ – Top $[_{\mathrm{TP}}$ – $[_{\mathrm{vP}}$ ... $[_{\mathrm{VP}}$ ... V] v] T]]]

Before turning to examples that illustrate this approach to selection and extraposition, let me address the second role played by [±wh] features that was mentioned above, viz., that of triggering *wh*-movement.

Interpretable, weak [±wh] on C can be checked by appropriate features on c-

[17]This approach presupposes that relative clause extraposition from NP is another type of process in German, triggered by a different factor. Indeed, there is independent evidence for formally differentiating between these two CP extraposition operations in German; see Kiss (1998).

[18]This is only possible if CP extraposition does not interrupt an TP-internal V cluster; cf. Haider (1994) and Büring & Hartmann (1997).

commanded XPs. The feature [+wh] on a *wh*-phrase behaves rather similarly to [+wh] on a matrix predicate. First, semantic considerations tell us that it is uninterpretable (but see Chomsky (1995, 290) for a different view): In a Karttunen-style semantics, *wh*-phrases are interpreted as simple existential quantifiers (cf. the above references), which also correlates with their morphological shape in a number of languages (cf. Cheng (1991), among many others). And second, since interpretable [±wh] on C is weak, it follows that uninterpretable [±wh] on a *wh*-phrase must be strong in German, so as to (initially) trigger *wh*-movement. However, there are two differences between [+wh] on a matrix predicate and [+wh] on a *wh*-phrase. For one thing, the latter feature must be obligatorily strong in German (*wh*-in situ is blocked in simple questions), whereas the former feature is only optionally strong (argument CP extraposition is optional). For another, strong [+wh] on a *wh*-phrase triggers (Greed-like) movement to a higher position, whereas strong [+wh] on a matrix predicate triggers movement by attraction.[19] As for instantiating uninterpretable, strong [–wh] on an XP, I will simply assume that this combination is systematically precluded in German, as a result of a feature co-occurrence restriction (cf. Gazdar, Klein, Pullum & Sag 1985) on numerations.[20]

From the above, it follows that [±wh] on C can play a dual role (for selection and *wh*-movement), in contrast to [±wh] on a matrix predicate, and [+wh] on a *wh*-phrase. This is so for one reason: [±wh] on C is interpretable and therefore

---

[19]Note that this approach is therefore incompatible with Chomsky's (1995) proposal that *all* movement is attraction. – Note also that, as it stands, this approach raises a problem for multiple questions, a property that it has in common with other minimalist approaches that locate the trigger for *wh*-movement on the *wh*-phrase, and not on C. (The alternative approach faces the reverse problem of accounting for multiple *wh*-movement as in, e.g., Bulgarian.) The problem is this: Given that [+wh] is strong on a *wh*-phrase, it seems that we should incorrectly expect (i-b) to be the correct form of a multiple question in German, and not (i-a).

(i) a. Wer_1 hat t_1 was_2 gesagt ?
who$_{nom}$ has what$_{acc}$ said
'Who said what?'

b. *Wer_1 was_2 hat t_1 t_2 gesagt ?
who$_{nom}$ what$_{acc}$ has said
'Who said what?'

There are several ways to solve this problem. E.g., there might be a meta-constraint on numerations that ensures weak [+wh] features of surplus *wh*-phrases in the case of multiple questions. Alternatively, one might follow Grewendorf (1997) in assuming that the surface representations in (i) are actually misleading, and that German does indeed have multiple overt *wh*-movement as in (i-b), but except for one, the heads of the *wh*-chains are realized as phonologically empty copies.

[20]There is no movement of non-*wh*-phrases to SpecC in German; cf.:

(i) *Ich denke [$_{CP}$ den Fritz_1 daß man t_1 fragen sollte]
I think ART Fritz$_{acc}$ that one ask should
'I think that one should ask Fritz.'

I will leave open how topicalization is to be accounted for, but, for the time being, it may suffice to state that here, different features are involved.

does not disappear after checking.[21] Table 1 summarizes the relevant assumptions about [±wh] features on *wh*-XP, C, and matrix predicates in German.

Table 1: [±wh] Features in German

Category	Value	Strength	Interpretability	Checking Position
XP	[+wh]	strong	not interpretable	–
XP	[–wh]	–	–	–
C	[+wh]	weak	interpretable	left-peripheral (wh-mvt)
C	[–wh]	weak	interpretable	–
V/v/T/Top	[+wh]	strong or weak	not interpretable	right-peripheral (extraposition)
V/v/T/Top	[–wh]	strong or weak	not interpretable	right-peripheral (extraposition)

The resulting system, while still fairly standard in nature, does not yet suffice to predict either the option of successive-cyclic *wh*-movement, or obligatory CP extraposition as a reflex of successive cyclicity. To this end, I will now adopt, as a central and non-trivial modification of the system in Chomsky (1995), the assumption that checking as required by LR takes place under identity of *features*, but not necessarily under identity of *feature values*, as is often (more or less tacitly) presupposed: Thus, a [–wh] feature on C can check the [+wh] feature of a *wh*-phrase.[22] However, such checking is clearly imperfect. It is a legitimate operation (due to feature identity), but nevertheless a failed attempt (due to a feature value clash). Suppose now that, as a result of imperfect checking, both features persist as such (whether interpretable or not); furthermore, since they are now both "activated," they are visible for overt syntax, i.e., strong. This can be formulated as in (17):

(17) *Feature Strengthening*:
Imperfect checking of two features αF, βF yields strong αF, βF.

This means that movement of a *wh*-phrase bearing the feature [+wh] to the specifier of a C head bearing the feature [–wh] is legitimate, but it will not remove any of the two features from the derivation. What is more, both features are now obligatorily strong, which implies that they have to be checked by two further movement operations, one applying to the *wh*-phrase (successive-cyclic movement), and one applying to C, or, rather, CP, by obligatory pied piping in German (extraposition). Before

[21] A question arises concerning [±wh] on root C nodes. For semantic reasons, we expect there to be a [+wh] feature on C in a root question; this feature can be checked by a *wh*-phrase (by an empty operator in yes/no questions, or perhaps by the finite verb in German, depending on the analysis one adopts for this kind of construction). The case is less clear with declarative root clauses. One possibility that strikes me as viable would be to assume that there is in fact no [–wh] feature on C present in this context. Alternatively, it might be tempting to identify a [–wh] feature on root C as the trigger for V/2 movement in German. Since the present approach does not force a decision here, I will leave the issue open.

[22] This implies that features that trigger movement are (at least) bivalent, as is standardly assumed for the feature [wh], and not monovalent, as recently proposed by Harley (1994), Ritter & Harley (1998) for a somewhat different class of features (those encoding person and number).

illustrating the effects of feature strengthening on successive-cyclic *wh*-movement and CP extraposition with some examples, let me briefly clarify three consequences that this revision has for the overall system developed in Chomsky (1995).

First, the operation of feature strengthening is to some extent similar to Watanabe's (1993) theory of derivational feature creation, according to which (Case-driven) movement to position π_1 creates a "secondary" feature that must then be checked off by further movement to a higher position π_2. In both approaches, the first movement step is legitimate as far as LR is concerned, but it automatically creates a need to follow this operation by a second one (which, in Watanabe's approach, is then successful, and which, in the approach suggested here, may or may not be successful, depending on whether checking is imperfect or not). As is well known, genuine derivational feature creation raises a problem vis-à-vis Chomsky's (1995, 228) Inclusiveness Condition according to which "no new objects are added in the course of computation." However, the present proposal evades this objection: feature strengthening may change the value of a given feature, but it does not introduce a new feature.

Second, it might be worth pointing out that imperfect checking is not the same as feature mismatch. The latter cancels the derivation (cf. Chomsky 1995, 308-309), which imperfect checking obivously must not do. It therefore becomes necessary to ensure that all standard cases of feature mismatch can be construed so as to involve different features, rather than identical features with different values. Consider, e.g., Case features. Suppose that we want to maintain that movement of an accusative NP to a position in which nominative Case is checked is blocked as a result of feature mismatch. Then, we are forced to adopt the view that the relevant features are Boolean in nature, i.e., bivalent [acc] and [nom], with values $\pm$. Crucially, we cannot assume a non-Boolean, multivalent [Case] feature (cf. [Case:acc], [Case:nom]). If the latter option were chosen, movement of an accusative NP to a nominative position would result in imperfect checking (and feature strengthening), and not in feature mismatch.[23]

Third and finally, table 1 suggests a correlation of strength and interpretability, such that only uninterpretable features can be strong. However, recall that imperfect checking of [$\pm$wh] features in the C domain results in a strong [–wh] feature on C, which, as noted above, is interpretable. Thus, minimally we have to conclude that if there is an incompatibility of strength and interpretability, it may hold in the numeration, but not necessarily throughout the derivation.[24]

With the theoretical framework laid out, let me turn to some empirical consequences for *wh*-movement and extraposition.

---

[23]At this point, various questions arise about successive-cyclic A-movement of [+nom]-marked NPs. Although the present approach might ultimately offer an interesting view on what triggers this operation, I will not pursue the matter here, mainly for reasons of space and coherence.

[24]That said, it might be that there is no interesting correlation in the first place. Cf. Brody (1995), Chomsky (1995), and Roberts & Roussou (1997) for discussion.

3.3. Consequences

3.3.1. German

A general prediction of the approach developed here is that movement operations that are otherwise optional (or confined to LF) are turned into obligatory ones in feature strengthening contexts. This is just what we find with argument CP extraposition in German.

Consider first the distribution of finite declarative (i.e., [–wh]) CP complements. This type of CP preferably shows up in post-verbal, extraposed position in German. However, as has often been noted (see Stechow & Sternefeld 1988, Webelhuth 1992, Büring 1995, and Bayer 1996; 1997, among many others), leaving it in a pre-verbal position is by no means impossible. Depending on a number of intervening factors (related, i.a., to intonation and choice of matrix predicate, on which see below), a pre-verbal finite declarative CP can be perceived as marked, as clumsy, as slightly deviant, or even as fully acceptable, but typically not as ungrammatical. This is shown in (18):

(18) a. daß er [$_{CP_2}$ daß Fritz Claudia liebt] gesagt/ nicht gewußt hat
that he that Fritz$_{nom}$ Claudia$_{acc}$ loves said/ not known has
'that he said/did not know that Fritz loves Claudia.'

b. daß er t_2 gesagt/ nicht gewußt hat [$_{CP_2}$ daß Fritz Claudia
that he said/ not known has that Fritz$_{nom}$ Claudia$_{acc}$
liebt]
loves
'that he said/did not know that Fritz loves Claudia.'

In the present approach, the example pair in (18) implies optional extraposition, and, in line with the standard line of reasoning in the minimalist program, we can assume that the apparant optionality of CP movement is to be reanalyzed as an obligatoriness of CP movement triggered by the optional presence of a strong [–wh] feature in the extedended projection (V, v, T, Top) of the matrix predicate. Thus, in (18-b), [–wh] on the matrix predicate is strong, thereby triggering CP extraposition, whereas in (18-a), [–wh] on the matrix predicate is weak, thereby delaying checking of the complement feature to LF. Note that since there is no overt checking operation in the domain of C_2 in either of the examples, there is no feature strengthening involved yet.

Next consider finite [+wh] CP complements. Again, extraposition is optional (although stylistically preferred).

(19) a. ?daß er [$_{CP_2}$ wen$_1$ Fritz t_1 liebt] nicht gesagt/ gewußt hat
that he whom$_{acc}$ Fritz$_{nom}$ loves not said/ known has
'that he did not say/know who Fritz loves.'

b. daß er t_2 nicht gesagt/ gewußt hat [$_{CP_2}$ wen$_1$ Fritz t_1 liebt]
that he not said/ known has whom$_{acc}$ Fritz$_{nom}$ loves
'that he did not say/know who Fritz loves.'

In analogy to declarative complements, we can assume that the [+wh] feature of the matrix predicate (V, v, T, Top) is optionally strong. Unlike what was the case in (18), there is overt checking in the domain of C_2: Since the [+wh] feature of the *wh*-phrase is, by assumption, strong in simple questions in German, movement of the *wh*-phrase to the specifier of C_2 is forced. However, since C_2 also bears a [+wh] feature (i.e., since the embedded CP acts as an indirect question, and the landing site of the *wh*-phrase corresponds to its LF scope position), this <[+wh],[+wh]> checking in the domain of C_2 qualifies as perfect. Hence, there is no feature strengthening in this case either, and CP extraposition stays optional, depending solely on the strength of the [+wh] feature on the matrix predicate.

Let me finally turn to finite [–wh] CPs from which *wh*-movement has occurred. As observed by Stechow & Sternefeld (1988), CP extraposition, which is optional otherwise (cf. (18)), becomes obligatory in this context:

(20) a. *(Ich weiß nicht) wen_1 er [$_{CP_2}$ t'_1 daß Fritz t_1 liebt] gesagt hat
I know not $whom_{acc}$ he that $Fritz_{nom}$ loves said has
'I do not know who he said that Fritz loves.'

b. (Ich weiß nicht) wen_1 er t_2 gesagt hat [$_{CP_2}$ t'_1 daß Fritz t_1
I know not $whom_{acc}$ he said has that $Fritz_{nom}$
liebt]
loves
'I do not know who he said that Fritz loves.'

This follows from the assumptions about imperfect checking and feature strengthening made above.[25] Irrespective of whether or not the [–wh] feature on the matrix predicate (V, v, T, Top) is strong, *wh*-movement to C_2 results in imperfect <[–wh],[+wh]> checking, and hence in strengthening of both the [+wh] feature of the *wh*-phrase and the [–wh] feature of CP_2. These two strong features must consequently be checked by two subsequent movement operations.

First, the strengthened [+wh] feature of NP_1 has to undergo movement to the next higher SpecC (pied piping the rest of NP_1). In the case at hand, this specifier position is in the domain of a C node that bears a [+wh] feature, so checking in that position is perfect, and nothing more needs to happen to either the *wh*-phrase or the embedded *wh*-clause (but of course, depending on the strength of the [+wh] selection feature of the root predicate, extraposition can in principle apply to the *wh*-clause – note that if it applies in (20-b), it does so string-vacuously). From a more general point of view, however, it is clear that the proposed mechanism also captures much longer *wh*-dependencies and systematically accounts for successive-

[25] I should mention that, in contrast to what is the case with the remnant movement data introduced in subsection 2.3., it would also be possible to account for the effect in (20) under the general assumption that V-CP order is base-generated in German, an assumption that I have abandoned (cf. subsection 2.3.2). On this alternative view, CP-V order as in (18-a) must be derived by CP scrambling in (20-a), and since CP_2 in (20-a) contains only bound traces, the same reasoning applies as with the freezing effect in (2-b) (cf. Webelhuth 1992).

cyclic *wh*-movement: If the surface position of NP_1 in (20-b) were a [–wh] position, imperfect checking and feature strengthening would apply anew, and the *wh*-phrase would have to move to the next higher SpecC position, and so on, until it reaches a SpecC position where perfect <[+wh],[+wh]> checking takes place.[26]

Second, the strengthened [–wh] feature of C_2 must be checked by a further syntactic operation. In this case, logical possibilities include (a) Merge in C_2 (affecting the shape of the complementizer); (b) Move to C_2 (V raising); (c) C_2 incorporation into the matrix predicate (i.e., minimal pied piping); and (d) CP_2 extraposition to the domain of the matrix predicate (i.e., maximal pied piping). While I see no reason to exclude any of these operations in principle (and at least some seem to be documented in one form or another in the world's languages), it is clear that the last strategy is the one that is chosen in German. Hence, if CP_2 extraposition applies after *wh*-extraction, as in (20-b), the strengthened [–wh] feature on C_2 is checked, and the derivation can converge; if, in contrast, CP_2 extraposition fails to apply after *wh*-extraction, as in (20-a), the strengthened [–wh] feature on C_2 remains unchecked in overt syntax, and the derivation crashes, yielding ungrammaticality. More generally, it follows that there will be as many obligatory CP extraposition operations as there are *wh*-movement steps to [–wh] SpecC positions in a derivation. This way, obligatory CP extraposition can be viewed as a genuine reflex of successive cyclicity, on a par with other, well-established reflexes of successive cyclicity in other languages (cf. the references in subsection 3.1).

A sample derivation of (20) is given in (21):

(21) a. *Pre-movement structure*:
$[_{CP}\ C_{[+wh]}$ er $[_{CP_2}\ [_C$ daß] Fritz $wen_{[+wh]}$ liebt] gesagt hat

b. *Wh-movement no. 1, feature strengthening*:
$[_{CP}\ C_{[+wh]}$ er $[_{CP_2}\ wen_{[+wh]}\ [_C$ daß $]_{[-wh]}$ Fritz t liebt] gesagt hat

c. *Wh-movement no. 2, no feature strengthening (= (20-a)):*
$[_{CP}$ wen $C_{[+wh]}$ er $[_{CP_2}\ t'\ [_C$ daß $]_{[-wh]}$ Fritz t liebt] gesagt hat]

d. *CP_2 extraposition to Top (= (20-b)):*
$[_{CP}$ wen $C_{[+wh]}$ er t_2 gesagt hat $[_{CP_2}\ t'_1\ [_C$ daß $]_{[-wh]}$ Fritz t liebt]]

Note that at this point, the assumption becomes relevant that the selection feature [±wh] can be instantiated on any head in the extended projection of V. If CP_2 extraposition in the last step in (21) were to target the domain of V, rather than that of Top, the highest head of V's extended projection, then SCC would be violated: The V domain is properly embedded in the C domain, and a strong [+wh] feature has already been checked in the C domain by the preceding *wh*-movement operation. Now, strictly speaking, the Top domain is also embedded in the C domain, so at first it appears as though the problem posed by SCC persists. However, Top is embedded

[26] A potential problem arises with respect to LR/MLC; it appears that the matrix predicate might be an unwanted intervening item blocking *wh*-movement from SpecC to SpecC, due to the fact that it has an unchecked [–wh] feature at the point at which *wh*-movement applies. For the moment, I will abstract away from this problem; I will return to this issue in subsection 4.2 below.

under C only minimally. It has been argued in Müller & Sternefeld (1993) that Top and C should be viewed as "competing" clausal functional heads that form two sides of the same coin, and that TopP and CP are "matching projections," in the sense that for many purposes, they act like a single projection. From this perspective, it makes sense to assume that for the purposes of SCC, the Top domain and the C domain count as one, so that SCC can be fulfilled by targeting the Top domain even if a preceding operation has targeted the immediately dominating C domain.[27]

3.3.2. Basque

As remarked above, there are logical possibilities other than obligatory CP extraposition to check off a strengthened [–wh] feature on an embedded C node. However, it is worth noting that the effect at hand is not confined to German. Basque appears to behave similarly to German with respect to *wh*-movement and CP extraposition, and it might therefore be amenable to the same analysis in terms of feature strengthening.[28]

Thus, extraposition of [–wh] CPs is in principle optional (cf. Ormazabal, Uriagereka, & Uribe-Etxebarria 1994).[29]

(22) a. Mirenek [$_{\text{VP}}$ [$_{\text{CP}_2}$ Parisera joango zela] esan] zidan
Mary$_{erg}$ to Paris go Aux-C say Aux
'Mary told me that she was going to Paris'

b. Mirenek [$_{\text{VP}}$ t_2 esan] zidan [$_{\text{CP}_2}$ Parisera joango zela]
Mary$_{erg}$ say Aux to Paris go Aux-C
'Mary told me that she was going to Paris'

Similarly, as noted by Ortiz de Urbina (1989), extraposition of [+wh] CPs is optional:

(23) a. [$_{\text{CP}_2}$ Noiz etorri d-en] galdetu dut
when come Aux-en ask Aux

[27] An alternative which I will not pursue here because it potentially creates unwanted ambiguities with respect to [wh] feature checking would be to assume that CP extrapostion in a derivation like (21) actually targets a CP-adjoined, rather than a TopP-adjoined position. This would imply that any *wh*-extraction from a finite CP in German creates an unbound intermediate trace, which would be in conflict with the generalization in (9) (though, incidentally, not with the actual account of the generalization given in this paper). See Müller (1998, ch. 6) for discussion.

[28] Basque and German are both SOV languages, and it is intriguing to pursue the idea that the choice of extraposition as a reflex of successive cyclicity is somehow related to this basic property. Hindi and Bengali are a priori interesting candidates to look at in this context, but in these languages, there are interfering factors. For one thing, Hindi and Bengali do not appear to exhibit overt *wh*-movement to SpecC (rather, they employ mechanisms like *wh*-in situ, *wh*-(long-distance) scrambling, and *wh*-scope marking); for another, extraposition is obligatory throughout (Hindi), or either obligatory or prohibited, depending on the type and position of the complementizer (Bengali). See Mahajan (1990), Davison (1991), Dayal (1991), Butt (1993) on Hindi, and Bayer (1996; 1997) on Bengali.

[29] In fact, it seems that leaving the CP in a pre-verbal position does not even create markedness.

'I have asked when he has come'

b. Ez t_2 dakit [$_{CP_2}$ noiz etorri d-en herri honetara]
neg know when come Aux-Sub this town
'I don't know when he has come to this town'

These facts follow if we assume that, as in German, the uninterpretable [±wh] selection feature of the matrix predicate is optionally strong in Basque, and that the interpretable [±wh] feature of the embedded C head is (initially) weak.

Crucially, Ormazabal, Uriagereka, & Uribe-Etxebarria (1994) observe that if *wh*-extraction takes place from a [–wh] CP, that CP is obligatorily extraposed:[30]

(24) a. *Zer$_1$ esan dizute [$_{CP_2}$ t'_1 irakurriko du-ela Peruk t_1] entzun
what$_{abs}$ say Aux read Aux-C Peru$_{erg}$ hear
dutela ?
Aux-C
'What did they tell you that they heard that Peter will read?'

b. Zer$_1$ esan dizute t_2 entzun dutela [$_{CP_2}$ t'_1 irakurriko du-ela
what$_{abs}$ say Aux hear Aux-C read Aux-C
Peruk t_1] ?
Peru$_{erg}$
'What did they tell you that they heard that Peter will read?'

Again, obligatory CP extraposition after *wh*-extraction can be viewed as a direct consequence of imperfect checking and feature strengthening in the embedded C system, and hence, as a reflex of successive cyclicity.

3.3.3. Partial Wh-Movement

Before returning to the problem of accounting for the ban on unbound intermediate traces in German, let me briefly point out another consequence of this approach to successive cyclicity. This consequence concerns the analysis of partial *wh*-movement constructions, which freely alternate with successive-cyclic *wh*-movement constructions in most varieties of German. A relevant pair of examples is given in (25):

(25) a. (Ich weiß nicht) was$_1$ er t_2 gesagt hat [$_{CP_2}$ wen$_1$ (daß) Fritz
I know not [+wh] he said has whom$_{acc}$ that Fritz$_{nom}$
t_1 liebt]
loves
'I do not know who he said that Fritz loves.'

b. (Ich weiß nicht) wen$_1$ er t_2 gesagt hat [$_{CP_2}$ t'_1 daß Fritz t_1
I know not whom$_{acc}$ he said has that Fritz$_{nom}$

[30] While the fundamental observation is due to Ormazabal, Uriagereka, & Uribe-Etxebarria (1994), the reasoning here is not: These authors analyze the phenomenon by assuming V-CP order to be base-generated, and CP-V order to be derived by leftward CP movement, in analogy to what has been proposed by Webelhuth (1992) for German; cf. footnote 25.

liebt]
loves
'I do not know who he said that Fritz loves.'

In the partial *wh*-movement construction (25-a), a *wh*-phrase fails to undergo overt movement to what seems to be its logical scope position – the SpecC position marked [+wh]; it stops in an intermediate SpecC position which bears a [–wh] feature (being embedded by a bridge verb which selects a declarative CP complement). The [+wh] target position is filled by a scope marker *was* ('what') (this scope marker may also show up in other intervening SpecC positions in longer *wh*-depedencies). In contrast, in the (truth-conditionally equivalent) standard long-distance *wh*-movement construction (25-b) (= (20-b)), the *wh*-phrase is moved to the SpecC position marked [+wh] in the syntax.

There are various competing analyses of the partial *wh*-movement construction in German, which can be grouped into different classes according to several characteristic features. Most notably, one can distinguish between "direct dependency" approaches (cf. van Riemsdijk 1982, Stechow & Sternefeld 1988, McDaniel 1989, Sabel 1996, and Müller 1997, among many others) and "indirect dependency approaches" (cf. Dayal 1994; 1996, Fanselow & Mahajan 1996, and Horvath 1997). Simplifying a bit, direct dependency approaches assume that there is some relation between the scope marker *was* and the partially moved *wh*-phrase, whereas indirect dependency approaches rely on the idea that there is a relation between the scope marker *was* and the *wh*-clause that immediately contains the moved *wh*-phrase.[31] However, it is a common property of *all* the analyses mentioned so far that they predict that there is no successive-cyclic overt *wh*-movement in (25-a) (in contrast to (25-b)). In other words: Neither of these approaches postulates an overt movement relation between the scope marker *was* and the partially moved *wh*-phrase.

The case is different with the version of the direct dependency approach developed by Cheng (1998). Based on the minimalist movement theory sketched in Chomsky (1995), Cheng (1998) suggests that (25-a) involves successive-cyclic *wh*-movement just as much as (25-b) does; the only difference is that in (25-b), the whole *wh*-phrase is pied piped twice by overt [+wh] feature movement, whereas in the partial movement construction (25-a), the *wh*-phrase is pied piped only once, and gets stranded in the intermediate [–wh] position by subsequent bare [+wh] feature movement. The moved [+wh] feature is phonologically realized by the scope

[31] There are further substantial differences among the approaches within each of the two classes. E.g., some of the direct dependency approaches postulate LF movement of the partially moved *wh*-phrase in addition to chain formation at S-structure (cf., e.g., Stechow & Sternefeld 1988, McDaniel 1989, Müller 1997), others do not (cf., e.g., van Riemsdijk 1982, Sabel 1996). Similarly, whereas Dayal's (1994; 1996) indirect dependency approach classifies scope marking *was* as a true argument that quantifies over propositions, and interprets the *wh*-clause as the restriction of this object, Fanselow & Mahajan (1996) and Horvath (1997) adopt the view that *was* is an expletive associated with the *wh*-clause, which in turn raises to the position of the expletive at LF. See the contributions in Lutz, Müller, & von Stechow (to appear) for further discussion of this and other issues pertaining to partial *wh*-movement and *wh*-scope marking.

marker *was*.[32]

Thus, we can classify approaches to partial *wh*-movement according to whether or not they postulate an overt movement relation between the scope marker *was* and the embedded *wh-phrase*, independently of the question of direct vs. indirect dependency. Recall now that the approach developed in subsection 3.2 above has identified obligatory CP extraposition as a reflex of successive-cyclic *wh*-movement. We are therefore in a position to use obligatory CP extraposition as a test for the adequacy of theories of partial *wh*-movement: If, on the one hand, a CP that immediately dominates a partially moved *wh*-phrase is treated by the grammar like an ordinary [–wh] CP (cf. (18)), or like an ordinary [+wh] CP in which perfect checking has taken place (cf. (19)), we expect that it does not have to extrapose. If, on the other hand, a CP that immediately dominates a partially moved *wh*-phrase is treated by the grammar like a [–wh] CP from which overt *wh*-movement has occurred (cf. (20)), we expect that it obligatorily undergoes extraposition.

As shown in (26) (vs. (25-a)), a CP in which partial *wh*-movement has taken place must undergo extraposition:

(26) *(Ich weiß nicht) was_1 er [$_{CP_2}$ wen_1 (daß) Fritz t_1 liebt] gesagt hat
I know not [+wh] he $whom_{acc}$ that $Fritz_{nom}$ loves said has
'I do not know who he said that Fritz loves.'

This supports an approach to partial *wh*-movement in German along the lines of Cheng (1998): Obligatory CP extraposition is a reflex of successive-cyclic overt *wh*-movement in German, and from the presence of this reflex, we may conclude that there is indeed an overt movement relation between the scope marker *was* and the position of the partially moved *wh*-phrase.[33]

If these considerations are on the right track, we also expect that the CP that immediately dominates the (highest) scope marker *was* does not have to undergo obligatory extraposition: Movement of the [+wh] feature of the *wh*-phrase, realized by *was*, to the target C domain bearing the feature [+wh] results in perfect check-

[32]This is actually a slight simplification of Cheng's (1998) approach, which relies on an articulated theory of pied piping, encoded in the notion of a repair strategy taking place after pure feature movement. For present purposes, though, the simplified version of Cheng's analysis may suffice.

[33]Of course, this argument is one of plausibility, and not logically conclusive. There might be other reasons that force obligatory extraposition in (26). However, it is not obvious what these reasons should look like. E.g., LF *wh*-movement does not have this effect on a [–wh] CP:

(i) a. ?(Ich weiß nicht) wem_1 sie t_1 [$_{CP_2}$ – daß Fritz wen_3 liebt] gesagt hat
I know not $whom_{dat}$ she that $Fritz_{nom}$ $whom_{acc}$ loves said has
'I do not know to whom she said that Fritz loves who.'

b. (Ich weiß nicht) wem_1 sie t_1 t_2 gesagt hat [$_{CP_2}$ – daß Fritz wen_3 liebt]
I know not $whom_{dat}$ she said has that $Fritz_{nom}$ $whom_{acc}$ loves
'I do not know to whom she said that Fritz loves who.'

The same goes for simple co-indexing relations: Replacing wen_3 in (i-ab) with a pronoun *sie* that is bound by the matrix subject does not create obligatory extraposition either.

ing. Hence, feature strengthening does not apply, and whether or not the CP is extraposed solely depends on the optional choice of strong vs. weak [+wh] selection feature of the matrix predicate. As shown in (27), this prediction is correct:

(27) a. ?daß ich [$_{CP_3}$ was$_1$ er t_2 gesagt hat [$_{CP_2}$ wen$_1$ Fritz t_1 liebt]]
that I [+wh] he said has whom$_{acc}$ Fritz$_{nom}$ loves
gar nicht wissen möchte
PRT not know want to
'that I really do not want to know who he said that Fritz loves.'

b. daß ich t_3 gar nicht wissen möchte [$_{CP}$ was$_1$ er t_2 gesagt hat [$_{CP_2}$
that I PRT not know want to [+wh] he said has
wen$_1$ Fritz t_1 liebt]]
whom$_{acc}$ Fritz$_{nom}$ loves
'that I really do not want to know who he said that Fritz loves.'

It is worth noting that the obligatory extraposition effect with partial *wh*-movement constructions is not confined to relatively short *wh*-dependencies, as in (25). Longer dependencies show exactly the same properties. Thus, by integrating Cheng's (1998) analysis of partial *wh*-movement in German into the approach to successive cyclicity developed here, we are led to conclude that the examples in (28) involve instances of imperfect <[–wh],[+wh]> checking in the C domains of CP_3 and CP_2, and an instance of perfect <[+wh],[+wh]> checking in the C domain of CP_4. Accordingly, both CP_3 and CP_2 undergo extraposition in (28-a), thereby checking the strengthened [–wh] features of the respective C heads. In contrast, in (28-b) CP_2 undergoes extraposition, but CP_3 fatally fails to do so, and ungrammaticality results due to an unchecked strong feature. Similarly, in (28-c), both CP_3 and CP_2 fail to undergo extraposition, yielding the same consequence. Finally, (28-d) shows that perfect checking in the domain of CP_4 does not induce feature strengthening. As we would expect, extraposition of CP_3 and CP_2 remains obligatory, but extraposition of CP_4 does not have to take place.

(28) a. (Ich weiß nicht) [$_{CP_4}$ was$_1$ sie t_3 gemeint hat [$_{CP_3}$ was$_1$ er t_2
I know not [+wh] she meant has [+wh] he
gesagt hat [$_{CP_2}$ wen$_1$ Fritz t_1 liebt]]]
said has whom$_{acc}$ Fritz$_{nom}$ loves
'I do not know who she thought that he said that Fritz loves.'

b. *(Ich weiß nicht) [$_{CP_4}$ was$_1$ sie [$_{CP_3}$ was$_1$ er t_2 gesagt hat [$_{CP_2}$
I know not [+wh] she [+wh] he said has
wen$_1$ Fritz t_1 liebt]] gemeint hat]
whom$_{acc}$ Fritz$_{nom}$ loves meant has
'I do not know who she thought that he said that Fritz loves.'

c. *(Ich weiß nicht) [$_{CP_4}$ was$_1$ sie [$_{CP_3}$ was$_1$ er [$_{CP_2}$ wen$_1$ Fritz
I know not [+wh] she [+wh] he whom$_{acc}$ Fritz$_{nom}$
t_1 liebt] gesagt hat] gemeint hat]
loves said has meant has
'I do not know who she thought that he said that Fritz loves.'

d. ?daß ich [$_{CP_4}$ was$_1$ sie t$_3$ gemeint hat [$_{CP_3}$ was$_1$ er t$_2$ gesagt hat
that I [+wh] she meant has [+wh] he said has
[$_{CP_2}$ wen$_1$ Fritz t$_1$ liebt]]] gar nicht wissen möchte
whom$_{acc}$ Fritz$_{nom}$ loves PRT not know want to
'that I really do not want to know who she thought that he said that Fritz loves.'

Thus, to end this detour, we have seen that partial *wh*-movement patterns with (20), rather than with (18)/(19), with respect to obligatory/optional CP extraposition. Under present assumptions, this suggests that imperfect checking is involved, which in turn lends support to the idea that partial *wh*-movement constructions involve overt scope marker movement, as proposed by Cheng (1998), in contrast to what is assumed in most of the literature on *wh*-scope marking constructions in German.

4. Remnant Movement, Successive-Cyclic Movement, Extraposition

With the approach to successive cyclicity developed in the last section as background, we can now return to the ban on unbound intermediate traces created by successive-cyclic *wh*-movement in German. The three relevant examples are repeated in (29) (where (29-a) = (12), (29-b) = (13), and (29-c) = (8-a)).

(29) a. *[$_{VP_3}$ t$_2$ Gesagt [$_{CP_2}$ t$'_1$ daß Fritz t$_1$ liebt]] weiß ich nicht [$_{CP}$ wen$_1$
said that Fritz$_{nom}$ loves know I not who$_{acc}$
er t$_3$ hat]
he has
'I do not know who he said that Fritz loves.'

b. *[$_{VP_3}$ [$_{CP_2}$ t$'_1$ Daß Fritz t$_1$ liebt] gesagt] weiß ich nicht [$_{CP}$ wen$_1$ er
that Fritz$_{nom}$ loves said know I not who$_{acc}$ he
t$_3$ hat]
has
'I do not know who he said that Fritz loves.'

c. *[$_{CP_2}$ t$'_1$ Daß Fritz t$_1$ liebt] weiß ich nicht [$_{CP}$ wen$_1$ er t$_2$ gesagt
that Fritz$_{nom}$ loves know I not whom$_{acc}$ he said
hat]
has
'I do not know who he said that Fritz loves.'

Recall from subsection 2.3.2 that unbound intermediate traces in V-CP topicalization structures as in (29-a) are ruled out as a freezing effect, by the interaction of BC and SCC (CP extraposition must follow *wh*-movement so as to respect BC, but the VP-adjoined landing site of CP extraposition is then too low to prevent a SCC violation). Thus, it remains to be shown that the prohibition against unbound intermediate traces in CP-V topicalization structures as in (29-b), and in CP topicalization structures as in (29-c), can also be derived from independently established constraints on movement, without construction-specific assumptions. This will be done in the following two subsections.

4.1. Remnant CP-V Topicalization

Turning to CP-V topicalization structures as in (29-b) first, it is clear that *wh*-movement of NP_1 to the domain of C_2 results in imperfect checking, hence in feature strengthening of both [+wh] on NP_1 and [–wh] on C_2. Strong [+wh] on NP_1 is subsequently checked by *wh*-movement to the [+wh] target position, but since CP_2 extraposition does not apply, a strong [–wh] feature remains at Spell-Out, and the derivation crashes, essentially for the same reason that (20-a) does.[34]

4.2. Remnant CP Topicalization

Finally, as far as bare remnant CP topicalization is concerned, note that the derivation of (29-c) must involve a combination of CP_2 extraposition and CP_2 topicalization: All alternative derivations violate BC, SCC, or the prohibition against strong features at Spell-Out. The only a priori remaining derivation that does not violate any of the constraints adopted so far is given in (30):[35]

(30) a. *Pre-movement structure*:
– weiß ich nicht [$_{CP_3}$ – er [$_{CP_2}$ – [$_{C_2}$ daß] Fritz wen_1 liebt] gesagt hat]

b. *Wh-movement of NP_1 to C_2, imperfect <[–wh],[+wh]> checking, feature strengthening*:
– weiß ich nicht [$_{CP_3}$ – er [$_{CP_2}$ wen_1 [$_{C_2}$ daß] Fritz t_1 liebt] gesagt hat]

c. *Wh-movement of NP_1 to C_3, perfect <[+wh],[+wh]> checking*:
– weiß ich nicht [$_{CP_3}$ wen_1 er [$_{CP_2}$ t'_1 [$_{C_2}$ daß] Fritz t_1 liebt] gesagt hat]

d. *CP_2 extraposition to Top, forced by strengthening of [–wh] on C_2*:
– weiß ich nicht [$_{CP_3}$ wen_1 er t_2 gesagt hat [$_{CP_2}$ t'_1 [$_{C_2}$ daß] Fritz t_1 liebt]]

---

[34] If CP_2 extraposition targets the V domain, a SCC violation occurs; cf. (29-a). Furthermore, the prediction arises that, ceteris paribus, a combination of remnant VP topicalization and a preceding CP extraposition operation should not be ruled out as strongly ungrammatical if CP extraposition targets a higher position, i.e., the Top domain. As shown in (i), this type of example does indeed represent an improvement over (29-a), (29-b).

(i) ??[$_{VP_3}$ t_2 Gesagt] weiß ich nicht [$_{CP}$ wen_1 er t_3 hat] [$_{CP_2}$ t'_1 daß Fritz t_1 liebt]
said know I not who_{acc} he has that $Fritz_{nom}$ loves
'I do not know who he said that Fritz loves.'

That said, it appears as though (i) is somewhat worse than comparable *wh*-island violations with remnant topicalization that do not involve successive-cyclic *wh*-movement; cf. (7). One might speculate that (i) raises additional parsing problems, due to multiple crossing of movement paths, but other than that, I will have to leave this issue open.

[35] Two remarks are due on the last step in (30). First, it is clear that CP topicalization must be triggered by a strong feature; however, as before, I will remain uncommitted here as to what the nature of this feature is. Second, there might be an additional (string-vacuous) step of CP_3 extraposition after CP_2 topicalization, depending on whether or not the matrix predicate has a strong [+wh] selection feature (but note that CP_3 extraposition is not forced by feature strengthening because CP_3 is a [+wh] clause in which perfect checking has occurred).

e. *CP_2 topicalization*:
* [$_{CP_2}$ t'_1 [$_{C_2}$ daß] Fritz t_1 liebt] weiß ich nicht [$_{CP_3}$ wen$_1$ er t_2 gesagt hat t'_2]

The question arises of what is wrong with this derivation. I would like to suggest that the problem is that CP_2 is first extraposed and then topicalized, which qualifies as an instance of improper movement.

Note first that, irrespective of the actual theoretical implementation of this idea, there is direct empirical evidence which shows that extraposed CPs cannot be topicalized. So far, I have confined myself to matrix predicates where finite CP extraposition is optional as such; but there are some predicates where this does not hold, i.e., where the embedded CP must be extraposed; cf. Webelhuth (1992), Büring (1995), Bayer (1996), and Müller (1996). One such predicate is *freuen* ('rejoice'):

(31) a. daß ich mich t_2 gefreut habe [$_{CP_2}$ daß Fritz gewonnen hat]
that I REFL rejoiced have that Fritz$_{nom}$ won has
'that I rejoiced in the fact that Fritz won.'
b. *daß ich mich [$_{CP_2}$ daß Fritz gewonnen hat] gefreut habe
that I REFL that Fritz$_{nom}$ won has rejoiced have
'that I rejoiced in the fact that Fritz won.'

For the time being, we may leave open the question of why exactly the choice of a weak [±wh] selection feature creates illformedness with these predicates, and simply note the fact as such.[36]

Interestingly, as noted in the references just given, these predicates strictly prohibit topicalization of their CP complements:

(32) *[$_{CP_2}$ Daß Fritz gewonnen hat] habe ich t_2 mich gefreut t'_2
that Fritz$_{nom}$ won has have I REFL rejoiced
'I rejoiced in the fact that Fritz won.'

From this we can conclude that the sequence extraposition → topicalization applying to a single item is impossible. In principle, it seems that one could leave it at that: The derivation in (30) involves a sequence of exactly this type, and therefore, we can conclude that (30) is ungrammatical for the same reason that (32) is. Still, it would of course be preferable if an account of this effect could be given. Indeed, the prohibition against extraposition feeding topicalization can be made to follow from various theories of improper movement. I will sketch two such theories, one that centers around a uniformity condition on chain formation, and one that is based on

[36] An ad hoc solution that might work for the cases at hand would be to assume that predicates of this type obligatorily bear a strong [±wh] selection feature, and do not permit a weak [±wh] feature in the first place. However, closer inspection reveals that this would be an over-simplification. In particular, it seems that with this kind of predicate, there is more structure than meets the eye in the embedded complement (an additional empty PP shell, i.e., a "hidden" preposition; cf. Büring (1995)). This additional structure might eventually be responsible for creating the obligatory extraposition effect. These considerations, however, do not directly bear on the issue at hand.

an economy constraint.

First, it has been argued by Müller & Sternefeld (1993; 1996), Takahashi (1994), and Collins (1997) that syntactic theory should recognize a more general notion of improper movement that the one adopted in Chomsky (1981), to the effect that chain formation cannot succeed if it involves different types of positions in a single A-bar chain.[37]

For concreteness, consider the Principle of Unambiguous Binding (PUB) developed in Müller & Sternefeld (1993; 1996):

(33) *Principle of Unambiguous Binding (PUB):*
An A-bar trace that is α-bound must be β-free.

The PUB blocks movement from one type of derived A-bar position to any other type of position; e.g., from a *wh*-position to an A-position (which accounts for illicit super-raising), or from an extraposition position to a *wh*-position (which eventually accounts for the fact that whereas most islands are strict for leftward movement, they are not strict for extraposition, cf. Müller (1996)), or from an extraposition position to a topic position. This last prediction is of course relevant for the case at hand. The extraposition trace t_2 in (30-e) is bound both by the intermediate trace t'_2 in the extraposed adjoined position, and by the chain antecedent CP_2 in the matrix topic specifier position, in violation of the PUB. (The same reasoning applies with the extraposition trace t_2 in (32).)

As an alternative to this kind of approach, one might try to dispense with a specific uniformity condition on chain formation, and derive improper movement effects by invoking an economy constraint. A relevant candidate is the transderivational economy constraint *Fewest Steps* that is used (in different formulations) by Epstein (1992), Collins (1994), Chomsky (1993; 1995), and Kitahara (1997), among others. Here, the basic idea would be that, from the point of view of economy, the CP extraposition operation in (30-d) is superfluous because it is subsequently undone by CP topicalization. More specifically, suppose that the economy constraint *Fewest Steps* (FS) forces the choice of the derivation with the smallest number of strong features: A strong feature is licensed only if it has an effect on PF output (cf. Chomsky 1995, 294 for a closely related proposal).

(34) *Fewest Steps (FS):*
Minimize the number of strong features required to reach a given PF string.

It seems that a constraint of this type can derive most of the PUB effects in Müller & Sternefeld (1993; 1996), and can thus replace a specific uniformity condition on chain formation in favour of something more general. More important in the present context, though, is the fact that FS blocks the derivation Δ_1 in (30) in favour of a

[37] In contrast, the theory of improper movement in Chomsky (1981), which relies on principle C of the binding theory, blocks mixing of different types of positions in a chain only in case successive-cyclic movement to an A-position takes place via an intermediate A-bar (SpecC) position, i.e., in case successive-cyclic super-raising applies.

competing derivation Δ_2 in which vacuous CP_2 extraposition does not take place: CP_2 extraposition in Δ_1 is subsequently undone by CP_2 topicalization, and therefore has no effect on PF output. Δ_2 in (35) avoids the strong [–wh] feature on C_2 by moving the *wh*-phrase to its target position C_3 in one step, thereby skipping over the C_2 domain; hence, no feature strengthening will apply. Given that CP is a barrier for any XP that it dominates unless XP is in SpecC, as has been argued by Baker (1988), Sportiche (1989), and Müller & Sternefeld (1993; 1995) (cf. footnote 4), this means that Δ_2 involves a violation of BC (CP_2 is a barrier for *wh*-movement that skips over the specifier of C_2), and is thus excluded. This more economical, fatally blocking derivation is given in (35):[38]

(35) a. *Pre-movement structure*:
– weiß ich nicht [$_{CP_3}$ – er [$_{CP_2}$ – [$_{C_2}$ daß] Fritz wen$_1$ liebt] gesagt hat]

b. *Wh-movement of NP_1 to C_3, perfect <[+wh],[+wh]> checking, BC violation*:
* – weiß ich nicht [$_{CP_3}$ wen$_1$ er [$_{CP_2}$ – [$_{C_2}$ daß] Fritz t$_1$ liebt] gesagt hat]

c. *CP_2 topicalization*:
* [$_{CP_2}$ – [$_{C_2}$ daß] Fritz t$_1$ liebt] weiß ich nicht [$_{CP_3}$ wen$_1$ er t$_2$ gesagt hat]

Thus, FS forces the choice of Δ_2 in (35) (which violates BC) over Δ_1 in (30) (which violates no local constraint – as we have seen, it respects BC, SCC). However, at this point a question arises that concerns the role of LR/MLC. Given the formulation in (14), it seems that, quite independently of whether or not CP_2 is a barrier for one-swoop *wh*-movement in (35), this one-swoop movement should violate MLC, hence LR that MLC is part of, by skipping over the specifier of C_2. Although this is indeed the prediction made by LR/MLC in its present formulation (14), and although this prediction seems to be welcome at first sight, it turns out that there is an independent reason for revising LR/MLC in such a way that it is not violated in (35), and that, therefore, recourse to a BC violation is necessary in (35).

The independent reason that necessitates a modification of LR/MLC has already been alluded to in footnote 26; it is related to the fact that the present system permits imperfect checking: It must be ensured that *wh*-movement from SpecC to SpecC

[38]Actually, there is a bit more to be said. CP_2 topicalization in (30) and (35) also crosses a CP barrier since it skips over the SpecC position, and the same goes for other examples involving topicalization from a *wh*-island which we have seen to exhibit an intermediate (subjacency-like) degree of illformedness, and not complete ungrammaticality (cf. (6), (7)). That this asymmetry with respect to CP barriers is one between topicalization and *wh*-movement in German, and not one between skipping over an empty SpecC position and skipping over a filled SpecC position, is suggested by *wh*-movement from a *wh*-island in German, which is far worse than topicalization in this context. As noted in subsection 2.3.1, I will not attempt to develop a full account of this phenomenon here; on the view adopted in Müller & Sternefeld (1993), the crucial difference is that *wh*-movement in this context has to violate BC twice (CP and TopP are barriers), whereas topicalization violates BC only once (only CP is a barrier).

is not blocked by intervening predicate heads with the feature [–wh]. As it stands, such a predicate head would incorrectly be predicted to disallow *wh*-movement from SpecC to SpecC – when *wh*-movement from a SpecC position to the next higher SpecC position occurs, an intervening predicate head must still have an unchecked [–wh] feature in German (had this [–wh] feature been checked by CP extraposition at this point, a BC violation would necessarily result with *wh*-extraction). The conclusion that must be drawn from this is that whereas checking as such continues to be possible under identity of features, and not necessarily under identity of feature values, so that imperfect checking is legitimate (cf. (36-a)), the MLC must now refer to identity of feature values, and not simply to identity of features (cf. (36-b)).

(36) *Last Resort (LR)* (modified):
α can raise to position β only if (a) and (b) hold:
a. α checks a feature F in β.
b. There is no γ with an unchecked feature F with an identical value that is closer to β (MLC).

As a result of the modified LR/MLC, *wh*-movement to the closest [–wh] SpecC position (on the way to the [+wh] SpecC target position) is permitted, but not forced by LR/MLC – such movement is forced by BC, because a CP is a barrier for everything bar its specifier.

To sum up, whatever account of the improper movement effect in (30) (uniformity or economy) eventually proves superior, it follows that the third and final of the a priori problematic constructions involving unbound intermediate traces listed in (29) is now also accounted for. Still, something remains to be said about ordinary CP topicalization, where no unbound traces are involved.

4.3. Non-Remnant CP Topicalization

The question arises of why simple non-remnant CP topicalization is possible, given that the CP_1's [–wh] feature must eventually be checked in a right-adjoined position by the matrix predicate in examples like (37) (= (5-a)).

(37) [$_{CP_1}$ Daß Fritz Claudia liebt] hat er t_1 gesagt (t'_1)
that Fritz$_{nom}$ Claudia$_{acc}$ loves has he said
'He said that Fritz loves Claudia.'

The problem is that CP_1 in (37) appears to be in a position at Spell-Out that is higher than the extraposition site in which clausal selection of CP_1 by the matrix predicate must be checked. Hence, at first sight it seems that (37) should involve improper movement, just like (30) does.

Interestingly, this problem is not confined to the approach developed here. A similar dilemma shows up with the interaction of *wh*-movement and Case-driven object raising in Chomsky (1995). The relevant construction is given in (38):

(38) [$_{CP}$ Who$_1$ did you [$_{vP}$ (t'_1) v [$_{VP}$ see t_1]]] ?

Here, the problem is not one of improper movement, but rather the fact that the feature involved in Case checking of objects must be weak in English (there is no overt object shift in this language). Therefore, it is unclear how the object *wh*-phrase who_1 in (38) could first move to the intermediate Specv position before reaching its target position, SpecC. Chomsky discusses a number of ways to reconcile constructions like (38) with the idea that every overt movement step must be triggered by a strong feature. One suggestion, which I will adopt here for the sake of concreteness, rests on the idea that t'_1 is not present at Spell-Out; movement to SpecC in (38) is in one swoop. Object Case checking is attained by LF raising of the trace t_1 to the intermediate Case position – an operation that is intuitively counter-cyclic but does not violate SCC in the formulation adopted here.[39]

From a general point of view, it is likely that whatever permits (38) will also permit (37), and this is indeed the case under the present suggestion: Just like Case checking of *wh*-object NPs in English, clausal type checking of topicalized clauses in German examples like (37) can be achieved by LF movement of the trace t_1 to the position of t'_1 (i.e., the only derivation that will succeed in generating the string in (37) is one in which the [–wh] feature of the matrix predicate is weak).

4.4. Conclusion

This concludes the discussion of the ban on unbound intermediate traces in finite CPs. All three crucial types of construction in (29) are now ruled out by independently established constraints of movement theory, and a recourse to construction-specific assumptions that explicitly refer to unbound traces or intermediate traces (or both) has proven unnecessary. Furthermore, the illformedness of the three constructions in (29) is reduced to what essentially amounts to a single common factor: Successive-cyclic *wh*-movement triggers obligatory CP extraposition in German, which, in remnant movement constructions, may then induce a SCC violation (with subsequent V-CP topicalization, as in (29-a)), or may result in an illicit strong feature at Spell-Out (with subsequent CP-V topicaliztion, as in (29-b)), or may induce improper movement (with subsequent CP topicalization, as in (29-c)).

---

[39]This makes it necessary to abandon another constraint envisaged in Chomsky (1995, 304) that precludes movement of traces. Furthermore, it should be noted that this solution, like its competitors, is not completely unproblematic. It depends on the assumption that traces can bear features that must be checked (which in turn seems to suggest a copy theory approach to movement, which has so far not proven necessary here), and it forces us to base checking theory not on individual items, but on chains (so that the fact that who_1 in (38) as such still has an unchecked Case feature is irrelevant as long as there is one chain member that has its Case feature checked). This latter stipulation, which hardly meets minimalist requirements, is a general consequence of the copy theory, but it is especially pressing in the case at hand because it is not the head of the chain on which the feature is checked, but an intermediate member; cf. Gärtner (1997, 122). For further discussion of constructions of this type and alternative proposals, see Bobaljik (1995), Groat & O'Neil (1996), Stabler (1996), and Gärtner (1997).

5. Infinitives

Control infinitives are interesting from the present perspective because their behaviour differs from that of finite CPs in a systematic manner, with respect to both *wh*-extraction and remnant movement.

5.1. Wh-Movement and Extraposition

Note first that, in contrast to what was the case with finite clauses, *wh*-movement from infinitives may occur without extraposition. Thus, both (39-a) and (39-b) are well formed:

(39) a. Was_1 hat der Fritz $[_{\alpha_2}$ (t'_1) $[_{\text{TP}}$ t_1 zu lesen]] abgelehnt ?
what$_{acc}$ has ART Fritz to read refused
'What did Fritz refuse to read?'

b. Was_1 hat der Fritz t_2 abgelehnt $[_{\alpha_2}$ (t'_1) $[_{\text{TP}}$ t_1 zu lesen]] ?
what$_{acc}$ has ART Fritz refused to read
'What did Fritz refuse to read?'

This is unexpected if there is imperfect <[–wh],[+wh]> checking within the infinitive α in the course of long-distance *wh*-movement: Imperfect checking triggers feature strengthening, hence, in German, obligatory extraposition. However, there is clear evidence that there is neither perfect <[+wh],[+wh]> checking, nor imperfect <[–wh],[+wh]> checking in German infinitives. As noted by Tappe (1984), there are no embedded *wh*-infinitives in German; cf. (40-a). Tappe's conclusion is that German infinitives simply lack the relevant structure that may provide the landing site for *wh*-movement; the minimal assumption within the approach adopted here is that German infinitives, whatever their categorial status, lack the [+wh] feature that may undergo perfect checking with the [+wh] feature of a *wh*-phrase.[40] In addition to this, (40-b) shows that there is no partial *wh*-movement in German infinitives. From this, we can conclude that infinitives also lack the feature [–wh], so that imperfect checking is not possible either.

---

[40]The full generality of the ban on *wh*-movement in German infinitives is sometimes questioned on the basis of apparent instances of clausal pied piping in relative clauses, as in (i).

(i) das Buch $[_{\text{CP}}$ das_1 $[_{\alpha}$ t_1 zu lesen] der Fritz beabsichtigt]
the book which to read ART Fritz intends
'the book that Fritz intends to read'

In van Riemsdijk's (1985) view, (i) involves true clausal pied piping, i.e., movement of a non-finite *wh*-CP that includes α to SpecC. In contrast, Haider (1985) and Grewendorf (1986) argue that only the relative pronoun has undergone *wh*-movement to SpecC in (i), and that α occupies a scrambling position to the right of C; under this second view, the ban on *wh*-movement in German infinitives can be maintained without qualification. The issue cannot yet count as settled; here, I will confine myself to pointing out one strong argument for the second approach: Varieties of German that permit a simultaneous filling of SpecC and C tolerate a relative complementizer *wo* ('where') to the left of α in (i), but not to the right of α.

(40) a. *Fritz weiß nicht [$_{\alpha_2}$ welches Buch$_1$ t$_1$ zu lesen]
Fritz$_{nom}$ knows not which book$_{acc}$ to read
'Fritz does not know which book to read.'

b. *Was$_1$ hat er versucht/abgelehnt [$_{\alpha_2}$ welches Buch$_1$ t$_1$ zu lesen] ?
[+wh] has he tried/refused which book$_{acc}$ to read
'Which book did he try/refuse to read?'

There are several possible explanations on the basis of the present approach. Either German infinitives are CPs, just like finite clauses are, the only difference being that infinitival C is not equipped with a [±wh] feature; or German infinitives are headed by some other category, which might then simply be T (more or less in line with Tappe's original proposal), or which might be some other functional category F which lacks a [±wh] feature. The difference between these options appears to be partly terminological (C vs. F, given that categories are to be viewed as sets of features, and F would differ from "standard" C only in lacking the feature [±wh]), and partly substantial, depending on matters which are not immediately relevant to the issues at hand and which I have nothing to say about here (e.g., the question of whether an impoverished TP structure for control infinitives would create problems for the theory of PRO).

For the sake of concreteness, I will assume that infinitives differ from finite clauses in German in being bare TPs; non-finite T differs from finite C in lacking a [±wh] feature.[41] However, it is clear that there is a selection relation between the matrix predicate and the infinitive. Following Fabb (1984) and, essentially, Bech (1955/57), I will adopt the view that this selection relation is mediated by a [V-Case] ("verbal Case") feature which is checked by the matrix predicate, by overt extraposition or LF raising, just as with finite clauses. Crucially, the [V-Case] feature on T cannot be checked with the [+wh] feature of a *wh*-phrase – in this case, a true feature mismatch would arise, and not just imperfect checking.

With respect to (39) and (40), this yields the following consequences. First, *wh*-movement to the α domain (now identified as TP) is correctly predicted to be impossible in (40): Neither perfect, nor imperfect checking of [±wh] is available in that position, for the simple reason that there is no head that could bear this feature. Second, it also follows that long-distance *wh*-movement from infinitives does not establish an intermediate trace t$'_1$ in (39).[42] Since there is no imperfect

[41] Nevertheless, I will continue to represent infinitives as α, so as to indicate that the other two options just envisaged would work equally well for the issues considered here.

[42] Since $\alpha \neq$ CP, no barrier is crossed – unlike CP, TP is not a barrier for *wh*-movement. Under the assumption that infinitives are CPs or FPs lacking a [±wh] feature, we would have to ensure that these XPs do not act as barriers for items that are not located in SpecC/SpecF. This implies that the notion of an XP barrier would have to (partly) depend on the properties of the head X, i.e., X's feature specification. This is accomplished in the theory of barriers developed in Müller & Sternefeld (1995), according to which an L-marked functional XP can be a barrier only if it licenses an A-bar specifier SpecX. As noted, this is not the case with infinitives in German, and therefore they would not qualify as barriers even if they exhibited an additional CP/FP structure.

checking here, feature strengthening cannot take place with the [V-Case] feature in the α domain. Hence, α extraposition is not forced: Whether or not infinitival extraposition takes place depends only on the choice of strong vs. weak [V-Case] feature on the matrix predicate.

5.2. Remnant Movement

Given that infinitives behave differently from finite clauses with respect to imperfect checking and feature strengthening, a number of interesting predictions arise for remnant infinitives in the three syntactic contexts in (29).

Let us begin with the context comparable to bare remnant CP topicalization (cf. (29-c), repeated once more in (41-a)). The relevant example involves remnant topicalization of an infinitive α across a *wh*-island created by the antecedent of the unbound trace. Recall that (41-a) is ruled out as an instance of improper movement, reducible to either a uniformity condition on chain formation (like the PUB), or to transderivational economy (FS), because it involves an illicit sequence of extraposition followed by topicalization. However, in contrast to finite CPs, *wh*-extraction from infinitives does not lead to feature strengthening. Hence, there is no obligatory extraposition after *wh*-extraction, and thus, no forced improper movement. We therefore expect that remnant infinitive topicalization should not be impossible. As shown in (41-b), this prediction is borne out: (41-b) is far from being perfect, but it does not seem to be worse than other instances of (complement) topicalization from *wh*-islands in German (cf. the examples in (6) and (7)), and it is at any rate much better than (41-a).

(41) a. *[$_{CP_2}$ t'_1 Daß Fritz t_1 liebt] weiß ich nicht [$_{CP}$ wen$_1$ er t_2 gesagt
that Fritz$_{nom}$ loves know I not whom$_{acc}$ he said
hat t'_2]
has
'I do not know who he said that Fritz loves.'

b. ?[$_{\alpha_2}$ Dem Fritz t_1 zu geben] weiß ich nicht [$_{CP}$ was$_1$ sie t_2
ART Fritz$_{dat}$ to give know I not what$_{acc}$ she
versucht/abgelehnt hat]
tried/refused has
'I do not know what she tried/refused to give to Fritz.'

Next, consider infinitival examples that are analogous to remnant CP-V topicalizations as in (29-b) (repeated here as (42-a)). Recall that (42-a) is ruled out because it involves a strengthened [–wh] feature on C_2 at Spell-Out that arose as a consequence of successive-cyclic *wh*-extraction from CP_2. Since *wh*-extraction from an infinitive does not involve imperfect checking, there is no strengthening of the [V-Case] feature of the infinitival head. Assuming that the [V-Case] feature of the matrix predicate is weak in the numeration underlying the derivation, we expect remnant α-V topicalization across a *wh*-island created by the antecedent of the unbound trace to be possible, in contrast to the analogous situation with finite CPs. Again, the prediction is correct (modulo the standard *wh*-island effect); cf. (42-b).

(42) a. *[$_{VP_3}$ [$_{CP_2}$ t'_1 Daß Fritz t_1 liebt] gesagt] weiß ich nicht [$_{CP}$ wen_1 er
that Fritz$_{nom}$ loves said know I not who$_{acc}$ he
t_3 hat]
has
'I do not know who he said that Fritz loves.'

b. ?[$_{VP_2}$ [$_{\alpha_3}$ Dem Fritz t_1 zu geben] versucht/abgelehnt] weiß ich
ART Fritz$_{dat}$ to give tried/refused know I
nicht [$_{CP}$ was_1 sie t_2 hat]
not what$_{acc}$ she has
'I do not know what she tried/refused to give to Fritz.'

Turning finally to infinitival equivalents of remnant V-CP topicalization structures as in (29-a) (cf. (43-a)), it emerges that this time, the approach developed here does not predict a difference between finite and non-finite clauses. Recall that (43-a) is blocked as a SCC violation – in order to meet BC, *wh*-extraction from CP_2 must precede CP_2 extraposition, and subsequent CP_2 extraposition to a VP-adjoined position therefore ends up in a domain that is dominated by the (perfect) checking domain of the *wh*-phrase. In this case, the reasoning applies to infinitives in the same way: Extraposition is not forced by a strengthened [–wh] feature (which is not present on α's head), but rather, by a strong [V-Case] feature on the matrix V, and this operation violates SCC. Thus, the expectation is that remnant V-α topicalization should be just as impossible as remnant V-CP topicalization. As shown in (43-b), this prediction is confirmed as well: (43-b) is perceived as distinctly worse than (42-b), and seems to be more on a par with (43-a).[43]

(43) a. *[$_{VP_3}$ t_2 Gesagt [$_{CP_2}$ t'_1 daß Fritz t_1 liebt]] weiß ich nicht [$_{CP}$ wen_1
said that Fritz$_{nom}$ loves know I not who$_{acc}$
er t_3 hat]
he has
'I do not know who he said that Fritz loves.'

b. ?*[$_{VP_2}$ t_3 Versucht/Abgelehnt] [$_{\alpha_3}$ dem Fritz t_1 zu geben] weiß ich
tried/refused ART Peter$_{dat}$ to give know I
nicht [$_{CP}$ was_1 er t_2 hat]
not what$_{acc}$ he has
'I do not know what she tried/refused to give to Fritz.'

---

[43] As noted by Daniel Büring (p.c.), examples of this type improve drastically if α-extraposition can be reanalyzed as V raising in a coherent construction; cf. (i).

(i) ?[$_{VP_2}$ [$_{\alpha_3}$ t_1 t_4] Versucht [zu lesen$_4$]] weiß ich nicht [$_{CP}$ was_1 sie t_2 hat]
tried to read know I now what$_{acc}$ she has
'I do not know what she tried to read.'

Here, there is no α extraposition operation present that might violate SCC. Rather, what moves is a single V, and *wh*-extraction can take place from the infinitive α (that is otherwise phonologically empty) in situ.

6. Conclusion

Let me summarize the main results of this paper. I have tried to argue that the strict ban on unbound intermediate traces in remnant movement constructions in German does not have to be accounted for by construction-specific assumptions, but can be made to follow from an elaborate minimalist theory of successive-cyclic *wh*-movement in German that is independently motivated.

The approach to successive cyclicity developed here relies on the idea that successive-cyclic *wh*-movement is triggered by a feature strengthening operation that results from imperfect <[–wh],[+wh]> checking in the C domain, and that creates a strong [+wh] feature on a *wh*-phrase which must eventually be removed by perfect <[+wh],[+wh]> checking in a [+wh] scope position. Feature strengthening that results from imperfect <[–wh],[+wh]> checking in the C domain also creates a strong (although interpretable) [–wh] feature on C that must subsequently be checked. One possibility, chosen in German and, apparently, Basque, is CP extraposition, an instance of checking of the clausal type of a CP complement by the matrix predicate, which is otherwise optional in these languages. Consequently, obligatory CP extraposition can be viewed as a reflex of successive cyclicity in German. Therefore, obligatory CP extraposition in partial *wh*-movement constructions suggests that imperfect checking is involved, i.e., that scope markers are realizations of moved [+wh] features.

This approach to *wh*-movement and extraposition could then be shown to derive the three core cases that instantiate the ban on unbound intermediate traces (cf. (29)): This ban reduces to either (i) SCC (in remnant V-CP topicalization contexts); (ii) a strong feature at Spell-Out (in remnant CP-V topicalization contexts); or (iii) improper movement – PUB or FS (in remnant CP topicalization contexts). Remnant infinitives could be shown to be immune against (ii) and (iii), but not against (i), because they do not participate in imperfect checking, an assumption that is independently corroborated by *wh*-movement options. Finally, on a more general note, I hope to have contributed to the general enterprise of showing that, once we adopt a strictly derivational model of grammar like the one developed in Chomsky (1995), we can maintain that there is really nothing peculiar about remnant movement – its properties follow from general constraints on movement, and there is no reason to believe that the syntax treats unbound traces differently from bound traces.

Still, it goes without saying that the above discussion has raised at least as many questions as it has answered. Some have emerged in the preceding pages; others I have been silent about. Let me mention three open problems that are particularly conspicuous.

The first problem concerns cross-linguistic variation. It seems reasonable to postulate that imperfect checking underlies successive cyclicity in general, and not only in some languages. However, in contrast to German and Basque, many languages, among them English, do not seem to exhibit any overt reflex of successive-cyclic *wh*-movement via [–wh] SpecC positions. In principle, there seem to be two possible solutions. It might be that feature strengthening is parametrized in such a way that

it may leave a [–wh] feature on C that has been checked by a [+wh] feature of a *wh*-phrase unaffected. On this view, in some languages imperfect checking would not create strength; rather, it would simply yield the same strength value for a [±wh] feature that it had before imperfect checking. Alternatively, one might argue that English does have feature strengthening resulting from imperfect checking after all; however, the strengthened [–wh] feature on C is checked by Merge with a complementizer that happens to be homophonous to the one that shows up in ordinary C domains where no imperfect checking has taken place (cf. Chomsky 1995, 267; note that this Merge operation would not violate SCC.)

Another aspect of the problem of cross-linguistic variation is that at least some of the well-established reflexes of successive cyclicity mentioned above do not show exactly the same distribution as the obligatory CP extraposition reflex highlighted here. For instance, the choice of complementizer form in Modern Irish (cf. McCloskey 1979) at first sight fits into the system developed here very well. Whether *aL* or *goN* is chosen in an embedded [–wh] SpecC position depends on whether *wh*-movement to that position has taken place in the derivation, and this can be interpreted as the result of a Merge operation triggered by strengthening of an otherwise weak [–wh] feature. However, the problem is that this reflex of successive cyclicity in Modern Irish is not confined to imperfect <[–wh],[+wh]> checking (like obligatory CP extraposition in German); it also shows up with perfect <[+wh],[+wh]> checking in the target position of *wh*-movement. Similar considerations apply in the case of the pronoun selection reflex in Ewe (cf. Collins 1994), which poses the additional problem that it is obligatory in perfect <[+wh],[+wh]> checking configurations, and optional in imperfect <[–wh],[+wh]> checking configurations. Furthermore, Baković (1998) notes that the V-to-C raising reflex in Spanish is subject to considerable dialectal variation that concerns both the nature of the *wh*-phrase and, in present terminology, the issue of perfect vs. imperfect checking. Thus, it seems that even though a substantial number of reflexes of successive cyclicity that show up in various languages follow under the system developed here, more must be said about variation, and about the fact that reflexes of successive cyclicity may show up with perfect <[+wh],[+wh]> checking in some languages. Still, the fact that at least some languages (Ewe, varieties of Spanish) distinguish between perfect and imperfect checking in some way (something that German and Basque do systematically) can be taken as evidence that the approach developed here based on feature strengthening does not have to be called into question; rather, the amount of cross-linguistic variation in this domain seems to suggest that what looks like a homogeneous reflex in Irish might in fact be due to two different triggers that happen to yield the same result. Clearly, more work is necessary here.

A second problem raised by the preceding remarks concerns topicalization and perhaps, more generally, other A-bar movement types that seem to apply successive-cyclically. Unlike *wh*-movement, topicalization creates a strict island for any type of movement in German (including topicalization itself; cf. Müller & Sternefeld (1993) and reference cited there for analyses of this topic island effect). Hence, examples involving illicit unbound intermediate topic traces of the type in (29) are straightforwardly excluded without recourse to the hypothesis that obligatory CP

extraposition resulting from imperfect checking is responsible; cf. (44-a), (44-b).

(44) a. *[$_{CP_2}$ Daß der Fritz die Maria$_1$ liebt] glaube ich [$_{CP_3}$ dem
that ART Fritz$_{nom}$ ART Maria$_{acc}$ loves believe I ART
Karl$_3$ sagte t$_3$ keiner t$_2$]
Karl$_{dat}$ said no-one$_{nom}$
'I think that no-one told Karl that Fritz loves Maria.'

b. *[$_{CP_2}$ t$'_1$ Daß der Fritz t$_1$ liebt] glaube ich [$_{CP_3}$ die Maria$_1$
that ART Fritz$_{nom}$ loves believe I ART Maria$_{acc}$
sagte keiner t$_2$]
said no-one$_{nom}$
'I think that no-one told Karl that Fritz loves Maria.'

However, as shown in (45-ab), it seems that successive-cyclic topicalization triggers obligatory extraposition, just like *wh*-movement in German.[44]

(45) a. ?Die Maria$_1$ habe ich t$_2$ gedacht [$_{CP_2}$ t$'_1$ daß der Fritz t$_1$ liebt]
ART Maria$_{acc}$ have I thought that ART Fritz$_{nom}$ loves
'I thought that Fritz loves Maria.'

b. *Die Maria$_1$ habe ich [$_{CP_2}$ t$'_1$ daß der Fritz t$_1$ liebt] gedacht
ART Maria$_{acc}$ have I that ART Fritz$_{nom}$ loves thought
'I thought that Fritz loves Maria.'

This suggests that imperfect checking (hence feature strengthening) is involved in the embedded C_2 domain in (45). This would follow if topicalization were a subcase of *wh*-movement, i.e., if it involved checking of a [+wh] feature. Although this is indeed sometimes assumed (cf. Chomsky 1977 and many others), I believe that, in view of the substantial number of asymmetries between topicalization and *wh*-movement in German (some of which have been mentioned above; cf. Müller & Sternefeld (1993) for an overview), this assumption is hardly tenable. I have avoided a precise characterization of topicalization in German in this paper, but, at least at first sight, it seems that the assumptions made here lead us to conclude that the feature triggering topicalization must be sufficiently unlike [+wh] to account for various asymmetries between the two movement types, and sufficiently similar to [+wh] to permit imperfect checking with [–wh] C nodes. A satisfying solution to this problem qrequires an elaborate analysis of the formal properties of topicalization. This, however, is beyond the scope of this paper.[45]

Finally, the question arises of what predictions are made for successive-cyclic *wh*-movement at LF, an issue that has been discussed controversially in the Principles and Parameters approach (cf., e.g., Chomsky 1981 vs. Lasnik & Saito 1984). Un-

[44]For reasons unclear to me, (45-a) is highly marginal in Northern varieties of German. The contrast to (45-b) is clear enough, though.

[45]An analogous situation arises with relativization in German, but here, the idea that [+wh] feature checking is involved strikes me as much less problematic.

der the analysis adopted here, successive-cyclic *wh*-movement at LF depends on the existence of a strengthened (or, at least, undeleted) [+wh] feature on a *wh*-phrase in [–wh] SpecC (and, in addition, on the irrelevance of BC – but not of LR/MLC – for [+wh] feature LF movement, given overt CP extraposition). For the approach developed in Chomsky (1995), it is crucial that all strong features must be checked by Spell-Out, and that strong features cannot be covertly inserted; however, it is not clear whether strong features that arise after Spell-Out as a result of imperfect checking and feature strengthening would pose a problem for that approach. If closer inspection should reveal that they do, one might revise feature strengthening accordingly for the LF component (such that imperfect checking leaves the uninterpretable feature intact, but does not actually strengthen it). Alternatively, there might be reasons to dispense with the notion of successive-cyclic *wh*-movement at LF altogether. Again, I will have to leave this question open.

7. References

Baker, Mark. 1988. *Incorporation.* Chicago: The University of Chicago Press.

Baković, Eric. 1998. Optimality and Inversion in Spanish. In *Is the Best Good Enough?*, eds. Pilar Barbosa et al., 35-58. Cambridge, Mass.: MIT Press & MITWPL.

Bayer, Josef. 1990. Notes on the ECP in English and German. *Groninger Arbeiten zur Germanistischen Linguistik* 30:1-51.

Bayer, Josef. 1993. V(P)-Topicalization and the Role of Traces. Ms., Universität Stuttgart.

Bayer, Josef. 1996. *Directionality and Logical Form. On the Scope of Focussing Particles and Wh-in-situ.* Dordrecht: Kluwer.

Bayer, Josef. 1997. CP-Extraposition as Argument Shift. In *Rightward Movement*, ed. Dorothee Berman, David LeBlanc, & Henk van Riemsdijk, 37-58. Amsterdam: Benjamins.

Bech, Gunnar. 1955/57. *Studien über das deutsche Verbum Infinitum.* Tübingen: Niemeyer (reprint 1983).

Beck, Sigrid. 1996. *Wh-Constructions and Transparent Logical Form.* Doctoral dissertation, Universität Tübingen.

Besten, Hans den, & Gert Webelhuth. 1987. Adjunction and Remnant Topicalization in the Germanic SOV-Languages. Paper presented at the GLOW-Conference, Venice, 1987.

Besten, Hans den, & Gert Webelhuth. 1990. Stranding. In *Scrambling and Barriers*, ed. Günther Grewendorf and Wolfgang Sternefeld, 77-92. Amsterdam: Benjamins.

Bobaljik, Jonathan. 1995. In Terms of Merge. *MITWPL* 27:65-103.

Brody, Michael. 1995. *Lexico-Logical Form.* Cambridge, Mass.: MIT Press.

Browning, M.A. 1989. ECP $\neq$ CED. *Linguistic Inquiry* 20:481-491.

Browning, M.A. 1991. Bounding Conditions on Representations. *Linguistic Inquiry* 22:541-562.

Büring, Daniel. 1995. On the Base Position of Embedded Clauses in German.

Linguistische Berichte 159:370-380.
Büring, Daniel, & Katharina Hartmann. 1997. Doing the Right Thing. *The Linguistic Review* 14:1-42.
Butt, Miriam. 1993. *The Structure of Complex Predicates in Urdu*. Doctoral dissertation, Stanford University.
Cheng, Lisa. 1991. *On the Typology of Wh-Questions*. Doctoral dissertation, MIT, Cambridge, Mass.
Cheng, Lisa. 1998. Moving just the Feature. Ms., University of California, Irvine. To appear in Lutz et al. (to appear).
Chomsky, Noam. 1977. On Wh-Movement. In *Formal Syntax*, ed. Peter Culicover, Tom Wasow, & Adrian Akmajian, 71-132. New York: Academic Press.
Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In *The View from Building 20*, eds. Kenneth Hale & Samuel Jay Keyser, 1-52. MIT, Cambridge, Mass.
Chomsky, Noam. 1995. Categories and Transformations. Chapter 4 of *The Minimalist Program*. Cambridge, Mass.: The MIT Press.
Chomsky, Noam. 1998. Minimalist Inquiries: The Framework (First version). Ms., MIT, Cambridge, Mass.
Cinque, Guglielmo. 1990. *Types of A'-Dependencies*. Cambridge, Mass.: MIT Press.
Clements, George, James McCloskey, Joan Maling, & Annie Zaenen. 1983. String-Vacuous Rule Application. *Linguistic Inquiry* 14:1-17.
Cole, Peter, & Gabriella Hermon. 1998. Partial Wh-Movement: Evidence from Malay. Ms., University of Delaware. To appear in Lutz et al. (to appear).
Collins, Chris. 1994. Economy of Derivation and the Generalized Proper Binding Condition. *Linguistic Inquiry* 25:45-61.
Collins, Chris. 1997. *Local Economy*. Cambridge, Mass.: MIT Press.
Davison, Alice. 1991. Finiteness and Case in Hindi-Urdu Complements. Ms., University of Iowa.
Dayal, Veneeta Srivastav. 1991. *Wh-Dependencies in Hindi and the Theory of Grammar*. Doctoral dissertation, Cornell University.
Dayal, Veneeta. 1994. Scope Marking as Indirect Wh-Dependency. *Natural Language Semantics* 2:137-170.
Dayal, Veneeta. 1996. Scope Marking: In Defense of Indirect Dependency. In *Papers on Wh-Scope Marking*, ed. Uli Lutz & Gereon Müller, 107-130. Tübingen: Arbeitspapiere des SFB 340 no. 76. Revised version in Lutz et al. (to appear).
Epstein, Samuel David. 1992. Derivational Constraints on $\overline{\text{A}}$-Chain Formation. *Linguistic Inquiry* 23:235-259.
Fabb, Nigel. 1984. *Syntactic Affixation*. Doctoral dissertation, MIT, Cambridge, Mass.
Fanselow, Gisbert. 1987. *Konfigurationalität*. Tübingen: Narr.
Fanselow, Gisbert. 1993. The Return of the Base Generators. Ms., Universität Passau. (Also in *Groninger Arbeiten zur Germanistischen Linguistik* 36:1-74.)
Fanselow, Gisbert, & Anoop Mahajan. 1996. Partial Movement and Successive

Cyclicity. In *Papers on Wh-Scope Marking*, ed. Uli Lutz & Gereon Müller, 131-161. Tübingen: Arbeitspapiere des SFB 340 no. 76. Revised version in Lutz et al. (to appear).
Ferguson, Scott, & Erich Groat. 1994. Defining "Shortest Move". Ms., Harvard University.
Fiengo, Robert. 1977. On Trace Theory. *Linguistic Inquiry* 8:35-61.
Gärtner, Hans-Martin. 1997. *Generalized Transformations and Beyond*. Doctoral dissertation, Universität Frankfurt/Main.
Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, & Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwell.
Grewendorf, Günther. 1986. Relativsätze im Deutschen. *Linguistische Berichte* 105:409-434.
Grewendorf, Günther. 1994. Variable Binding and Reconstruction. Ms., Universität Frankfurt/Main. (*Sprachwissenschaft in Frankfurt*, Report no. 12.)
Grewendorf, Günther. 1997. Anti-Superiority and Multiple Wh-Fronting. Ms., Universität Frankfurt/Main.
Grewendorf, Günther, & Joachim Sabel. 1994. Long Scrambling and Incorporation. *Linguistic Inquiry* 25:263-308.
Grimshaw, Jane. 1991. Extended Projection. Ms., Rutgers University.
Groat, Eric, & Wayne O'Neil. 1996. Spell-Out at the Interface. In *Minimal Ideas*, ed. Werner Abraham, Samuel Epstein, Höskuldur Thrainsson, & Jan-Wouter Zwart, 113-139. Amsterdam: Benjamins.
Haider, Hubert. 1985. Der Rattenfängerei muß ein Ende gemacht werden. Ms., Universität Wien.
Haider, Hubert. 1991. Fakultativ kohärente Infinitivkonstruktionen im Deutschen. Ms., Universität Stuttgart. (Arbeitspapiere des SFB 340 no. 17.)
Haider, Hubert. 1994. Detached Clauses – The Later the Deeper. Ms., Universität Stuttgart. (Arbeitspapiere des SFB 340 no. 41.)
Harley, Heidi. 1994. Hug a Tree: Deriving the Morphosyntactic Feature Hierarchy. *MITWPL* 21:289-320.
Heim, Irene. 1992. Remarks on Wh-Questions. Ms., MIT, Cambridge, Mass.
Horvath, Julia. 1997. The Status of 'Wh-Expletives' and the Partial Wh-Movement Construction of Hungarian. *Natural Language and Linguistic Theory* 15:509-572.
Huang, Cheng-Teh James. 1993. Reconstruction and the Structure of VP: Some Theoretical Consequences. *Linguistic Inquiry* 24:103-138.
Karttunen, Lauri. 1977. Syntax and Semantics of Questions. *Linguistics and Philosophy* 1:3-44.
Kayne, Richard. 1998. Overt vs. Covert Movement. Ms., New York University.
Kiss, Tibor. 1998. The Semantics of Relative Clause Extraposition. Ms., IBM Heidelberg.
Kitahara, Hisatsugu. 1994. A Minimalist Analysis of Cross-Linguistically Variant CED Phenomena. *NELS* 24:241-253.
Kitahara, Hisatsugu. 1997. *Elementary Operations and Optimal Derivations*. Cambridge, Mass.: MIT Press.
Koster, Jan. 1987. *Domains and Dynasties*. Dordrecht: Foris.

Lasnik, Howard, & Mamoru Saito. 1984. On the Nature of Proper Government. *Linguistic Inquiry* 15:235-289.

Lutz, Uli, Gereon Müller, & Arnim von Stechow (eds). to appear. *Wh-Scope Marking*. Amsterdam: Benjamins.

McCloskey, James. 1979. *Transformational Syntax and Model Theoretic Semantics. A Case Study in Modern Irish*. Dordrecht: Reidel.

McDaniel, Dana. 1989. Partial and Multiple Wh-Movement. *Natural Language and Linguistic Theory* 7:565-604.

Mahajan, Anoop. 1990. *The A/A-bar Distinction and Movement Theory*. Doctoral dissertation, MIT, Cambridge, Mass.

May, Robert. 1977. *The Grammar of Quantification*. Doctoral dissertation, MIT, Cambridge, Mass.

Müller, Gereon. 1996. On Extraposition and Successive Cyclicity. In *On Extraction and Extraposition in German*, ed. Uli Lutz & Jürgen Pafel, 213-243. Amsterdam: Benjamins.

Müller, Gereon. 1997. Partial Wh-Movement and Optimality Theory. *The Linguistic Review* 14:249-305.

Müller, Gereon. 1998. *Incomplete Category Fronting*. Dordrecht: Kluwer.

Müller, Gereon, & Wolfgang Sternefeld. 1993. Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24:461-507.

Müller, Gereon, & Wolfgang Sternefeld. 1995. Extraction, Lexical Variation, and the Theory of Barriers. In *Lexical Knowledge in the Organization of Language*, ed. Urs Egli, Peter Pause, Christoph Schwarze et al., 35-80. Amsterdam: Benjamins.

Müller, Gereon, & Wolfgang Sternefeld. 1996. A-bar Chain Formation and Economy of Derivation. *Linguistic Inquiry* 27:480-511.

Ormazabal, Javier, Juan Uriagereka, & Maria Uribe-Etxebarria. 1994. Word Order and Wh-Movement: Towards a Parametric Account. Ms., University of Connecticut & University of Maryland.

Ortiz de Urbina, Jon. 1989. *Parameters in the Grammar of Basque*. Dordrecht: Foris.

Riemsdijk, Henk van. 1982. Correspondence Effects and the Empty Category Principle. *Tilburg Papers in Language and Literature* 12, University of Tilburg.

Riemsdijk, Henk van. 1985. On Pied-Piped Infinitives in German Relative Clauses. In *Studies in German Grammar*, ed. Jindřich Toman, 165-192. Dordrecht: Foris.

Ritter, Elizabeth, & Heidi Harley. 1998. Sorting Out You, Me, and the Rest of the World. A Feature-Geometric Analysis of Person and Number. Ms., University of Calgary & University of Pennsylvania.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.

Roberts, Ian, & Anna Roussou. 1997. Interface Interpretation. Ms., Universität Stuttgart & University of Bangor.

Sabel, Joachim. 1995. *Restrukturierung und Lokalität. Universelle Beschränkungen für Wortstellungsvarianten*. Doctoral dissertation, Universität Frankfurt/Main. Appeared 1996: Berlin: Akademieverlag.

Sabel, Joachim. 1996. Asymmetries in Partial Wh-Movement. In *Papers on Wh-Scope Marking*, ed. Uli Lutz & Gereon Müller, 289-315. Tübingen: Arbeitspapiere des SFB 340 no. 76. Revised version in Lutz et al. (to appear).

Sportiche, Dominique. 1989. Le Mouvement Syntaxique: Contraintes et Paramtres. *Langages* 95:35-80.

Stabler, Ed. 1996. *Acquiring and Parsing Languages with Movement.* Book ms., UCLA.

Stechow, Arnim von. 1996. Against LF Pied-Piping. *Natural Language Semantics* 4:57-110.

Stechow, Arnim von, & Wolfgang Sternefeld. 1988. *Bausteine syntaktischen Wissens.* Opladen: Westdeutscher Verlag.

Stowell, Tim. 1981. *Origins of Phrase Structure.* Doctoral dissertation, MIT, Cambridge, Mass.

Takahashi, Daiko. 1994. *Minimality of Movement.* Doctoral dissertation, University of Connecticut.

Tappe, Thilo. 1984. On Infinitival Clauses Without COMP. In *Sentential Complementation*, ed. Wim de Geest & Y. Putseys, 227-237. Dordrecht: Foris.

Thiersch, Craig. 1985. VP and Scrambling in the German Mittelfeld. Ms., University of Tilburg.

Torrego, Esther. 1984. On Inversion in Spanish and Some of Its Effects. *Linguistic Inquiry* 15:103-129.

Watanabe, Akira. 1993. *Agr-Based Case Theory and Its Interaction with the A-bar System.* Doctoral dissertation, MIT, Cambridge, Mass.

Webelhuth, Gert. 1992. *Principles and Parameters of Syntactic Saturation.* Oxford: Oxford University Press.

Zwart, Jan-Wouter. 1993. *Dutch Syntax. A Minimalist Approach.* Doctoral dissertation, Rijksuniversiteit Groningen.