Gereon Müller

Universität Tübingen, Germany

Incomplete Category Fronting

A Derivational Approach to Remnant Movement in German

Kluwer Academic Publishers

Dordrecht / Boston / London

1998

Contents

Preface

This monograph is concerned with constructions in German like those in (1):

(1) a. Gelesen hat das Buch keiner
read has the book$_{acc}$ no-one$_{nom}$

b. Zu lesen hat das Buch keiner versucht
to read has the book$_{acc}$ no-one$_{nom}$ tried

c. Zu lesen versucht hat keiner das Buch
to read tried has no-one$_{nom}$ the book$_{acc}$

d. Ein Buch hat Antje über die Liebe gelesen
a book$_{acc}$ has Antje$_{nom}$ about the love read

In all these cases, fronting of an “incomplete” category has taken place (a verbal category in (1-a), (1-b), and (1-c), and a nominal category in (1-d)). These items are incomplete in the sense that an internal argument of the head of the topicalized category has remained behind in the middle field of the clause. Thus, the examples in (1) differ minimally from those in (2), where the fronted categories are “complete”:

(2) a. Das Buch gelesen hat keiner
the book$_{acc}$ read has no-one$_{nom}$

b. Das Buch zu lesen hat keiner versucht
the book$_{acc}$ to read has no-one$_{nom}$ tried

c. Das Buch zu lesen versucht hat keiner
the book$_{acc}$ to read tried has no-one

d. Ein Buch über die Liebe hat Antje gelesen
a book$_{acc}$ about the love has Antje$_{nom}$ read

In these cases, we can assume that a full XP has been fronted – a VP in (2-a) and (2-c), an infinitival clause in (2-b), and an NP in (2-d). In seminal (though largely still unpublished) work carried out within the Principles and Parameters framework, Craig Thiersch (1985) and Hans den Besten & Gert Webelhuth (1987) have proposed that cases of incomplete category fronting as in (1) are amenable to basically the same kind of analysis as their counterparts in (2). According to this view, topicalization in (1-a) through (1-d) involves XP movement. However, the fronted XPs contain traces of the internal arguments that have been extracted prior to fronting; they are “remnant” XPs.

The operation that is responsible for this removal of an argument from a remnant XP is identified as Ross's (1967) scrambling rule: This movement rule transports the internal argument to a middle-field internal position before topicalization applies. Thus, in this approach the sentences in (1) all have the general format in (3):

(3) $[_{\mathrm{XP}}$ ... t_1 ... $]_2$ ... YP_1 ... t_2 ...

The remnant movement approach to incomplete category fronting that is embodied in (3) makes a number of interesting and welcome predictions, but, as is well known, it also raises a lot of problems. Four of these strike me as particularly relevant, and I want to address them in this study. First, it is evident that the remnant movement approach presupposes that traces can occur unbound: t_1 in (3) is not c-commanded by its antecedent; however, such a configuration seems to be illegitimate otherwise (due to the Proper Binding Condition). Second, the construction in (3) does not exhibit a Freezing effect: Moved items are usually islands and cannot contain traces whose antecedent is located outside of the moved item, but this Freezing generalization does not seem to hold in (3) (i.e., an Anti-Freezing effect shows up). Third, otherwise undocumented asymmetries between movement types seem to arise that are not predicted under the remnant movement approach per se; for instance, whereas (3) can be a grammatical structure if XP is topicalized, as in (1), it has often been noted that (3) is normally impossible if XP is scrambled; compare, for instance, (1-b) with (4):

(4) *daß zu lesen das Buch keiner versucht hat
that to read the book_{acc} no-one_{nom} tried has

Fourth and finally, it seems that there are restrictions on the kind of trace that can show up unbound in (3). In particular, intermediate traces that arise as a result of successive-cyclic movement seem to be barred in this construction. Thus, whereas the first two problems can be viewed as part of the more general question of why remnant movement exists at all, the latter two problems relate to the question of why remnant movement appears to behave differently from other, well-established instances of movement.

My main goal here is to account for these problems posed by the remnant movement approach to incomplete category fronting in German and to do so without invoking construction-specific assumptions (this implies, without special constraints that mention unbound traces). The main claim that I would like to argue for is the following: What at first sight looks like a peculiar property of remnant movement actually turns out not to be peculiar at all if we adopt a derivational theory of movement along the lines of Chomsky's minimalist program, which has grown out of the Principles and Parameters approach. The minimalist program is renowned for its initial lack of independent empirical motivation and for the fact that it forces us to reconsider and alter many successful analyses of syntactic phenomena that were developed within the more representational Principles and Parameters theory. However, I believe that in the domain of incomplete category fronting at least, the case is different: A strictly derivational movement theory that incorporates some of the main premises of the minimalist program offers the key to the solution of a number of recalcitrant problems raised by this construction.

Of course, although I think that such a derivational approach unravels some of the mysteries of incomplete category fronting and remnant movement, it is quite clear that it does not automatically solve all problems with remnant movement; in fact, many new problems arise in domains where one would not expect them to under a more representational approach. Let me name just one that is especially striking and conceptual in nature: In a derivational theory of movement, constraints are typically also derivational, i.e., they do not act as surface filters (in contrast to, e.g., most of the constraints in standard Principles and Parameters theory). Hence, it often does not suffice to look simply at the surface string of a sentence to determine whether or not a given constraint is violated; rather, complete derivations must be examined. This usually complicates syntactic argumentation significantly and thereby sometimes makes explanations look less simple than they do in a representational approach that relies on surface filters. In spite of this, though, I think that more is gained by adopting the derivational approach to incomplete category fronting than is lost by largely dispensing with surface filters, but this is, of course, up to the reader to decide for her/himself.

A few remarks are due concerning the scope of the present study. Despite its length, this monograph has an essay-like character, in the sense that I focus on a few problems and systematically neglect "fringe issues" that arise in the course of the discussion and would eventually have to be settled. Thus, the central rationale is not to give as comprehensive an overview of constructions involving incomplete category fronting as possible, but rather to adopt one specific approach and explore it in depth. In line with this, I will not be concerned with certain other cases that involve (or at least have been argued to involve) incomplete category fronting, such as the so-called "NP split" construction, as shown in (5):

(5) a. Bücher hat Karl keine gelesen
books$_{acc}$ has Karl$_{nom}$ none read
'As for books, Karl has not read any.'

b. daß Bücher Karl keine gelesen hat
that books$_{acc}$ Karl$_{nom}$ none read has

Despite some claims to the contrary that can be found in the literature, I believe that this construction differs substantially from the one exemplified by the data in (1).[1] Hence, omitting NP split constructions from an approach to incomplete category fron-

[1] I will name just a few differences: First, unlike cases of incomplete category fronting as in (1), NP split constructions show non-matching effects in the sense that the fronted item does not necessarily correspond to the gap postulated in the middle field. Thus, van Riemsdijk (1989, 106) gives sentences like (i-a), and Fanselow (1993, 59) even marginally accepts sentences like (i-b):

ting of the type in (1) seems justifiable to me.

Another case where the limited scope of the present study becomes apparent concerns cross-linguistic variation. Throughout this study, the main focus is on incomplete category fronting in German, but the proposed analyses tend to be formulated in a much more general way so that they make predictions for other languages, too. Accordingly, I sometimes adduce data involving incomplete category fronting in a number of other languages (among them Breton, Danish, Dutch, English, French, Italian, Japanese, Korean, Serbo-Croatian, and Swedish) and show how they relate to the analysis suggested for German. By and large, it seems to me that the predictions of the proposed system are more often than not confirmed by evidence from other languages, but it should be noted that I do not attempt to account for incomplete category fronting in other languages in a systematic way.

The book is organized as follows. In chapter 1, I introduce the phenomenon of incomplete category fronting in German. I present the two fundamentally different approaches to the construction (viz., in terms of remnant movement and in terms of movement of a "small" category), report some initial evidence in favour of the remnant movement approach, and then go on to develop the four problems related to this approach which I will strive to solve in the remainder of the book.

Since these solutions crucially rest on the adoption of (a specific version of) the derivational movement theory assumed in the minimalist program, I then leave remnant movement completely aside for a chapter and address the issue of movement theory in the minimalist program in chapter 2. At the end of this chapter, a particular combination of derivational locality constraints (Barriers Condition), economy constraints (Fewest Steps, Last Resort), and cyclicity constraints (Strict Cycle Condition) is adopted that serves as the basis for the approach to remnant movement that is carried out

---

(i) a. Einen Wagen hat er sich noch keinen leisten können
a car$_{acc}$ has he REFL yet none$_{acc}$ afford could
'As for cars, he has not been able to afford one yet.'

b. ?Raubvögel mag er nur Bussarde
birds of prey$_{acc}$ likes he only buzzards$_{acc}$
'As for birds of prey, he only likes buzzards.'

Second, the issue of whether NP split constructions involve movement at all or should be analyzed in terms of base-generation cannot yet count as settled. However, even if we assume movement to be involved, it seems highly unlikely that we are dealing with configurations that take the form of (3), i.e., that involve an unbound trace in the fronted item – but if there is no unbound trace involved, the two problems related to the Proper Binding Condition and Freezing effects do not arise in the first place. Third, as Grewendorf (1995, 1298) observes, NP split constructions are not sensitive to the kind of restriction that is documented in (4); e.g., as shown in (5-b), the fronted item in this construction can occupy a scrambling position, provided that an appropriate intonation pattern is present. Given these properties (and others that can be found in the literature, see, e.g., Fanselow (1988; 1993), van Riemsdijk (1989), Tappe (1989), Grewendorf (1995), and references cited in these articles), I think it is fair to conclude that the NP split construction in German is of a different type than the kind of incomplete category fronting I am interested in here and obeys (partially) different laws.

in the four ensuing chapters. In these chapters, the four fundamental problems with remnant movement are addressed in turn.

Chapter 3 is concerned with the problem raised by the existence of unbound traces in light of the Proper Binding Condition. Here I show that the Proper Binding Condition can (in fact, should) be dispensed with in any form; illicit lowering is excluded by the Strict Cycle Condition, which, however, does not block remnant movement.

Chapters 4 and 5 form the core of this monograph. In chapter 4, I address the problem of Freezing and Anti-Freezing effects, and I show how the interaction of the Barriers Condition, Fewest Steps, Last Resort, and the Strict Cycle Condition yields the result that moved XPs normally are islands for extraction, whereas they are not in the case of remnant movement. Furthermore, the analysis is extended to extraposition, and it is shown that the well-known Anti-Freezing effect with extraposition in German can be explained by assuming (certain kinds of) extraposition to involve remnant movement, i.e., unbound traces.

Next, chapter 5 discusses the (apparent) sensitivity of remnant movement to the type of movement involved (e.g., remnant topicalization is possible, remnant scrambling is not). I argue that such a bifurcation cannot cover all data and that the correct underlying generalization is that the remnant XP and the antecedent of the unbound trace in a structure like (3) must occupy different positions; i.e., the two items cannot be affected by the same type of movement. This is first accounted for by introducing a (non-construction-specific) constraint of "Unambiguous Domination;" then, I attempt to derive the generalization from an independently motivated revision of the Last Resort condition that incorporates the Minimal Link Condition.

In chapter 6 I turn to the problem raised by ill-formed sentences that involve unbound intermediate traces in remnant movement constructions. I argue that it is neither necessary nor possible to maintain that there is a specific prohibition against unbound intermediate traces, and I try to show that the basic restrictions fall out of the theory developed so far, once certain facts about extraposition options in German are taken into account.

Finally, chapter 7 draws a conclusion and briefly sketches the consequences of this study for an inherently derivational approach to syntax.

Acknowledgements

Many people have contributed to this monograph over the last years, by offering helpful comments, raising interesting questions, providing data (or data judgements), etc. For all this, I would like to thank Steve Abney, Josef Bayer, Sigrid Beck, Kirsten Brock, Michael Brody, Daniel Büring, Chris Collins, João Costa, Danny Fox, Werner Frey, Hans-Martin Gärtner, Jochen Geilfuß, Willi Geuder, Günther Grewendorf, Jane Grimshaw, Hubert Haider, Tilman Höhle, Hanneke van Hoof, Kyle Johnson, Jung-Goo Kang, Shin-Sook Kim, Hisatsugu Kitahara, Hap Kolb, Uli Lutz, Anoop Mahajan, Kimberly Miller, Marga Reis, Bernhard Rohrbacher, Joachim Sabel, Ken Safir, Matthias Schlesewsky, Uli Sauerland, Peter Staudacher, Arnim von Stechow, Craig Thiersch, Hubert Truckenbrodt, Sten Vikner, Gert Webelhuth, and Chris Wilder. A special thank you goes to Gisbert Fanselow and Wolfgang Sternefeld, whose influence on this book is tremendous and who also provided detailed written comments. I also gratefully acknowledge the opportunity to present part of the material in this study to audiences at the 19th GGS meeting in Zürich (1993), the 9th Comparative Germanic Syntax Workshop at Harvard University (1994), the Potsdam University Workshop on Phrase Structure and Lexical Structure (1994), the LOT Winter School at Tilburg University (1995), the SFB-Workshop "Nachfeldsyntax" at Tübingen University (1995), the Utrecht University Workshop on Optionality (1995), the colloquium of the Graduiertenkolleg at Stuttgart University (1996), and a seminar at Potsdam University (1996/1997). Finally, I would like to thank the two reviewers for *Studies in Natural Language and Linguistic Theory* for their careful reading and their helpful comments, which led to numerous improvements.

Research for this monograph was supported by a DFG grant to the *Sonderforschungsbereich 340*, project A7 (Arnim von Stechow & Wolfgang Sternefeld, principal investigators). Part of the material in this book is also contained (in a somewhat different, earlier version) in articles that have appeared elsewhere. First, the analysis given in section 6 of chapter 2 is taken from the paper "Optional Movement and the Interaction of Economy Constraints" which has appeared in the volume *The Role of Economy Principles in Linguistic Theory*, edited by Manfred Bierwisch, Hans-Martin Gärtner, and Chris Wilder (Berlin: Akademieverlag, 1997, pp. 115-145). Second, parts of chapter 4 show up as the article "Extraposition as Remnant Movement" that appears in the volume *Rightward Movement*, edited by Dorothee Beermann, David LeBlanc, and Henk van Riemsdijk (Amsterdam: Benjamins, 1997). Finally, an earlier version of parts of chapter 5 is contained in the paper “A Constraint on Remnant Movement” that has appeared in *Natural Language and Linguistic Theory* 14 (1996), pp. 355-407.

Chapter 1

Incomplete Category Fronting

1. The Phenomenon

It is a well-known characteristic of German syntax that “incomplete” or “partial” constituents can be fronted. Consider the following examples involving topicalization of a verbal category.[1]

(1) a. [$_\beta$ Gelesen]$_1$ hat [$_{\mathrm{IP}}$ das Buch keiner t$_1$]
read has the book$_{acc}$ no-one$_{nom}$

b. [$_\beta$ Gelesen]$_1$ hat [$_{\mathrm{IP}}$ keiner das Buch t$_1$]
read has no-one$_{nom}$ the book$_{acc}$

In these examples, a participle has been moved in front of the finite verb in V/2 position. The central question connected with this movement is the following: Does the fronted verbal category β include a trace of the direct object (*das Buch* (‘the book’) in (1-a) and (1-b), or does β only contain a verb? This question has been answered controversially in the literature within the Principles and Parameters framework as developed in, for instance, Chomsky (1981; 1986) and Chomsky & Lasnik (1993).

On the one hand, Thiersch (1985) and den Besten & Webelhuth (1987; 1990) (and, following them, Stechow & Sternefeld (1988, ch.12), Huang (1993), Bayer (1993), and Grewendorf & Sabel (1994)) have argued that β in (1) is a full VP containing a trace of the direct object. What is fronted under this approach is a *remnant* category, i.e., a VP including a trace of the direct object that is not bound at S-structure. In this analysis, the derivation of (1-a) and (1-b) necessarily involves two instances of movement. First, the direct object *das Buch* is scrambled out of the VP and adjoined to IP in front of the subject in (1-a) and (string-vacuously) to VP in (1-b); second, the

[1] See, e.g., Paul (1919, par. 68) for some early discussion. Paul gives examples like *Gemahnt hab ich ihn schon oft* (lit. ‘Reminded have I him already often’), and notes: “In besonderen Fällen kann der Inf. oder das Part. wie ein anderes Satzglied als psychologisches Subj. an die Spitze des Satzes treten.”

remnant VP is topicalized. This is shown for (1-a) in (2):[2]

(2) $[_{\text{VP}}$ t_2 Gelesen $]_1$ hat $[_{\text{IP}}$ $[_{\text{NP}}$ das Buch $]_2$ $[_{\text{IP}}$ keiner t_1]]
read has the book_{acc} no-one_{nom}

The structure of (1-b), under this view, looks as in (3):[3]

(3) $[_{\text{VP}}$ t_2 Gelesen $]_1$ hat $[_{\text{IP}}$ keiner $[_{\text{VP}}$ $[_{\text{NP}}$ das Buch $]_2$ t_1]]
read has no-one_{nom} the book_{acc}

In what follows, I will refer to this analysis as the remnant movement approach.

On the other hand, it has been suggested by Fanselow (1983; 1992), van Riemsdijk (1989), and Frey & Tappe (1991), among others, that incomplete category fronting as in (1-a) does not involve remnant movement in the sense of (2) and (3); rather, β in (1-a) is assumed to be a complete verbal category that does not contain a trace of the direct object. This kind of analysis comes in two subvarieties. Fanselow (1983) and van Riemsdijk (1989) propose that verbal projections other than VPs can undergo topicalization in German (i.e., V' categories and V^0 categories as well); the structure of, e.g., (1-a) then looks as in (4-a).[4] Frey & Tappe (1991) and Fanselow (1992), in contrast, argue that although what has undergone movement in the sentences in (1-a) is a VP, the fronted category does not contain an unbound trace – in this approach, arguments of the verb are base-generated in VP-adjoined positions, so that a lower VP segment that does not dominate an argument trace can be topicalized; cf. (4-b):

(4) a. $[_{\text{V}}$ Gelesen $]_1$ hat $[_{\text{IP}}$ das Buch keiner t_1]
read has the book_{acc} no-one_{nom}

b. $[_{\text{VP}}$ Gelesen $]_1$ hat $[_{\text{IP}}$ $[_{\text{NP}}$ das Buch $]_2$ keiner $[_{\text{VP}}$ t_2 t_1]]
read has the book_{acc} no-one_{nom}

Clearly, the two approaches to incomplete category fronting in German (remnant movement vs. movement of a "small" verbal category) differ substantially, from a con-

---

[2] Huang (1993) and Bayer (1993) in addition assume that β also contains a trace of the subject, which follows under the Predicate-Internal Subject Hypothesis (cf. Kitagawa (1986), Sportiche (1988), Speas (1990), and Koopman & Sportiche (1991), among many others). Here and henceforth, I will mostly ignore possible VP-internal subject traces, unless this affects an issue under discussion; see in particular chapter 5.

[3] In (3), string-vacuous scrambling of NP_2 ends up in a VP-adjoined position, and remnant VP topicalization affects only the lower VP segment. Such an analysis might be problematic under several theories of projection and adjunction; see, e.g., Kayne (1994) and Kolb (1995). Alternatively, one might postulate (under the remnant movement approach) that NP_2 is adjoined to some functional XP between VP and IP. In what follows, I will abstract away from this issue, assuming adjunction to VP and topicalization of a VP-segment in (3), mainly for reasons of simplicity.

[4] This is only partially true in the case of van Riemsdijk (1989, 129), who suggests that after V- or V'-movement, the fronted category is subject to a regeneration operation that creates a full VP.

ceptual point of view as well as from an empirical one.[5] In the next section, I first give a well-known conceptual argument supporting the hypothesis that incomplete category fronting involves remnant movement in the sense of (2) and (3), and then go on to explore several empirical predictions that arise under this approach and that will be shown to lend support to it.

2. The Remnant Movement Approach

2.1. A Conceptual Argument for Remnant Movement

Given the general nature of the rule Move α, it is necessary to impose severe restrictions on its application and output. One such restriction recognized in Chomsky (1986) is this:

(5) Only XPs and X^0 categories can undergo movement.

From (5), it follows that the fronted items β in (1-a) are either XPs or X^0s. Whereas (5) does not yet force a decision between the competing approaches on the basis of examples like (1-a), it does already pose a problem for the analysis proposed by Fanselow (1983) and van Riemsdijk (1989) on the basis of (6).

(6) [$_\beta$ Ein Buch gegeben]$_1$ hat die Claudia dem Peter t_1
a book$_{acc}$ given has ART Claudia$_{nom}$ ART Peter$_{dat}$

Here, one internal argument of the verb goes along with it under topicalization (viz., the accusative NP *ein Buch* ('a book')), and the other internal argument (the dative NP *dem Peter*) stays behind. Clearly, then, under Fanselow's (1983) and van Riemsdijk's (1989) proposal, what has undergone topicalization must be a V' category (it cannot be a V^0 category, and a VP would automatically contain a trace), in violation of (5). Next, consider the standard assumption of X-bar theory in (7), put forward in Chomsky (1981; 1986) and Stowell (1981).[6]

(7) Every non-head is an XP.

---

[5]Interestingly, work on incomplete category fronting in German which has been carried out within the framework of HPSG reflects more or less the same fundamental dichotomy. Extraction in HPSG is handled in terms of the creation and passing on of a SLASH feature; so, Hinrichs & Nakazawa's (1994, 11ff) approach, which makes crucial use of a SLASH feature on an incomplete topicalized VP, is in many respects similar to the Principles and Parameters approach in terms of remnant movement. On the other hand, it is suggested by Nerbonne (1994, 112ff) and others that an incomplete topicalized VP does not bear a SLASH feature signalling the absence of one or more arguments of V; this approach is similar in certain respects to the Principles and Parameters approach that relies on the movement of "small" verbal categories.

[6]For an attempt to derive both the assumptions in (5) and (7) from independently established principles of movement and projection, cf. Chomsky (1995).

Given (7), even the data in (1-a) cannot be derived under the analysis of Fanselow (1983) and van Riemsdijk (1989) anymore – the landing site of topicalization, being a non-head position, can only be filled by an XP. However, both (5) and (7) are compatible with the analysis developed by Frey & Tappe (1991) and Fanselow (1992), according to which an example like, e.g., (6) looks as in (8) before topicalization applies:

(8) – hat die Claudia [$_{\mathrm{VP}}$ dem Peter [$_{\mathrm{VP}}$ ein Buch [$_{\mathrm{VP}}$ gegeben]]$_1$]
has ART Claudia$_{nom}$ ART Peter$_{dat}$ a book$_{acc}$ given

It is assumed that the indirect object and the direct object are base-generated in VP-adjoined positions, and VP-topicalization can then affect either the smallest VP (giving a result comparable to the sentences in (1-a)), or the immediately dominating VP containing the direct object and the verb (as in (6)), or the next higher VP that also includes the indirect object, as in (9).

(9) [$_{\mathrm{VP}}$ Dem Peter [$_{\mathrm{VP}}$ ein Buch [$_{\mathrm{VP}}$ gegeben]]]$_2$ hat die Claudia t$_2$
ART Peter$_{dat}$ a book$_{acc}$ given has ART Claudia$_{nom}$

Still, it seems that there are problems with Frey & Tappe's (1991) and Fanselow's (1992) approach. First, their approach necessitates a special assumption about the projection of verbs in German – it must be stipulated that arguments of V are never projected VP-internally, in contrast to what is the case in other languages. This special assumption conflicts with standard X-bar theory (see Stowell (1981), Chomsky (1981)) and with recent ideas developed in Chomsky (1995). Second and more importantly, it seems that this analysis has difficulty accounting for sentences like those in (10) in a straightforward way.

(10) a. [$_\beta$ Dem Peter gegeben]$_2$ hat die Claudia ein Buch t$_2$
ART Peter$_{dat}$ given has ART Claudia$_{nom}$ a book$_{acc}$

b. [$_\beta$ Einem Mann geschenkt]$_2$ hat sie die Kette t$_2$ nicht
a man$_{dat}$ given has she$_{nom}$ the chain$_{acc}$ not

Here, the indirect object, bearing dative Case, undergoes topicalization along with the verb (and not the direct object, as in (6)). To derive examples like those in (10) based on the analysis in Frey & Tappe (1991) and Fanselow (1992), it appears that one either has to relax the laws governing projection in German (such that, as an alternative to (8), a direct object can be base-generated above a dative-bearing indirect object in double object constructions in German), or one has to acknowledge the existence of an unbound trace of the direct object in the fronted VP in (10) after all.[7]

[7] Although there can be little doubt that the sentences in (10) are well formed from a purely grammatical point of view, markedness is also an issue here – sentences of the type in (10) (where the indirect object and the verb are fronted together, and the direct object remains in an IP-internal position) are generally perceived as more marked than sentences of the type in (6) (where the direct object and the verb are fronted together,

Similarly, the (relative) wellformedness of the examples in (11) poses a problem for the approach developed by Frey & Tappe (1991):

(11) a. $[_{\text{VP}}$ t_1 Vermutet $]_2$ habe ich sie $[_{\text{PP}}$ in Prag $]_1$ t_2
assumed have I her in Prague
'I have assumed that she is in Prague.'

b. $[_{\text{VP}}$ (t_2) t_1 Angestrichen $]_3$ hat er $[_{\text{NP}}$ das Haus $]_2$ [AP blau $]_1$ t_3
painted has he the house$_{acc}$ blue

Under Frey & Tappe's (1991, 9ff) assumptions, the non-argumental categories PP_1 and AP_1 in (11-a) and (11-b), respectively, are base-generated VP-internally, in contrast to arguments (like NP_2 in (11-b)). This implies that unbound traces show up in the fronted VPs, even under their assumptions – unless one is willing to adopt the view that neither arguments nor non-arguments (as in (11)) can show up VP-internally in German.[8]

Admittedly, these conceptual considerations are not yet compelling. If there is strong counter-evidence to the existence of unbound traces which the remnant movement approach developed by Thiersch and den Besten & Webelhuth presupposes, one might rightfully reconsider and call into question the overall framework which is embodied in assumptions like (5) and (7), and in the standard assumptions concerning lexical projection. Still, if such counter-evidence is absent, or can be explained away, it seems that there is every reason to adopt the remnant movement approach to incomplete category fronting, and not an approach along the lines of small category fronting, as envisaged by Fanselow, van Riemsdijk, and Frey & Tappe. Thus, in the remainder of this monograph, I will postulate that the remnant movement approach is in fact correct. In the following subsection, I address a welcome consequence of the remnant movement approach with respect to cross-linguistic variation.

2.2. Cross-linguistic Variation

As shown in (12), VP topicalization is an option in English:

(12) $[_{\text{VP}}$ Kicked the dog $]_2$ John never has t_2

---

and the indirect object stays behind). One might hold that these markedness considerations somehow relate to the issue of whether VP fronting retains 'normal' or 'unmarked' word order; but, for lack of an articulated theory of markedness as opposed to (un-) grammaticality, I will not pursue the matter any further here.

[8]It should be noted that examples like these are considered ungrammatical by Frey & Tappe. Similarly, Sabel (1995, 67-72) gives analogous examples that he considers ill formed, and he concludes that non-argument traces may never occur unbound. I do not follow these judgements here (or, in the case of some of Sabel's examples, I surmise that the illformedness has an independent reason), and assume that remnant movement may create both unbound argument traces and unbound non-argument traces, as long as no independent constraints are violated. That said, the issue of markedness as opposed to (un-) grammaticality that was brought up in the last footnote arises here, too.

However, unlike German, English does not exhibit incomplete category fronting as in (1-a); compare the ungrammatical English sentences in (13) to their German counterparts in (14) (= (1-a)).

(13) a. *[$_{\text{VP}}$ t$_2$ Kicked]$_1$ John [$_{\text{NP}}$ the dog]$_2$ never has t$_1$
b. *[$_{\text{VP}}$ t$_2$ Kicked]$_1$ John never has [$_{\text{NP}}$ the dog]$_2$ t$_1$

(14) a. [$_{\text{VP}}$ t$_2$ Gelesen]$_1$ hat [$_{\text{IP}}$ [$_{\text{NP}}$ das Buch]$_2$ keiner t$_1$]
read has the book$_{acc}$ no-one$_{nom}$
b. [$_{\text{VP}}$ t$_2$ Gelesen]$_1$ hat [$_{\text{IP}}$ keiner [$_{\text{NP}}$ das Buch]$_2$ t$_1$]
read has no-one$_{nom}$ the book$_{acc}$

Thiersch (1985) and den Besten & Webelhuth (1987; 1990) have observed that, according to the remnant movement approach, this instance of cross-linguistic variation follows without any further assumptions from the fact that German has a scrambling rule and English does not. Recall that under this approach, VP topicalization in examples like (14-a) and (14-b) is necessarily preceded by a scrambling operation that moves NP$_2$ (*das Buch* ('the book')) out of the VP. Clearly, one can predict that if a language does not exhibit scrambling, it will not exhibit remnant VP topicalization of this type either. English, in contrast to German, does not exhibit scrambling, and therefore the absence of constructions like those in (13) is explained straightforwardly: It is not remnant VP topicalization per se that is ungrammatical in (13); rather, the preceding scrambling operation which moves NP$_2$ (*the dog*) out of the VP is not permitted. Thus, the ungrammaticality of, e.g., (13-a), is derived in exactly the same way as the ungrammaticality of (15):[9]

(15) *John has [$_{\text{NP}}$ the dog]$_1$ never kicked t$_1$

In contrast, according to an approach in which incomplete category fronting does not involve unbound traces, it is far from clear why the English examples in (13) are ungrammatical.

The remnant movement approach is further corroborated by the observation that although English lacks constructions of the type in (1-a), there is no *general* prohibition against incomplete VP topicalization in English. Consider the following sentence (see Thiersch (1985, 13), Barss (1986, 409), Browning (1987, 362ff), Saito (1989, 199), Höhle (1991, 170f), and Huang (1993, 105ff), among others, for extensive discussion of this type of construction):

(16) [$_{\text{VP}}$ Criticized t$_2$ by his boss]$_1$ John$_2$ has never been t$_1$

---

[9]Of course, the question arises of why English lacks scrambling in the first place. Here, I will not have anything new to say about that issue; see, e.g., Koster (1986), Webelhuth (1988; 1992), Fanselow (1992; 1993; 1997), Haider (1993), Reuland & Kosmeijer (1993), Müller & Sternefeld (1993), and Müller (1995) for proposals.

In (16), a remnant VP has undergone topicalization, but in contrast to the examples in (13), ungrammaticality does not result. There is indeed a crucial difference between the ill-formed examples in (13-a) and (13-b), and the well-formed example in (16) – the topicalized VP headed by the passivized verb *criticized* in (16) contains a trace of Case-driven A-movement, rather than a trace left by scrambling. Given that A-movement to SpecI is available in English, the wellformedness of (16) is to be expected under the remnant movement approach.[10]

A look at other Germanic languages provides further evidence for reducing the contrast between the German examples in (14) and the English examples in (13) to the independently attested property of exhibiting or lacking a scrambling rule that generates structures with free word order. It appears that precisely those languages which lack scrambling fail to exhibit incomplete category fronting of the type in (1-a). This holds, e.g., for Swedish and Danish.[11] Dutch, however, which has scrambling, permits incomplete category fronting of the type in (1-a).[12]

---

[10]A certain type of construction is sometimes taken to show that remnant movement in English may in fact not "strand" antecedents in A-positions; cf. Saito (1989, 199), Lasnik & Saito (1992, 140ff), and literature cited there. A relevant example is (i), which is ungrammatical, in contrast to (ii):

(i) *[How likely t_1 to be a riot] is $there_1$?

(ii) $There_1$ is likely [t_1 to be a riot]

I will assume (following Höhle (1991), Huang (1993), and Takano (1993)) that examples like (i), and not examples of the kind in (16), represent the exception, so that we can ignore (i) for the present purposes.

[11]While these languages do not exhibit scrambling, they do not lack IP-internal extraction from VP completely: If (and only if) the main verb has moved out of VP, weak object pronouns obligatorily undergo raising out of the VP ("object shift") to a position following the finite verb (cf. Holmberg (1986) and Vikner (1994), among many others). Interestingly, at first sight it looks as though object shift can lead to remnant VP topicalization in Swedish, just as scrambling in German does (cf. Holmberg (1997)):

(i) $[_{VP}$ Kysst t_1 $]_2$ har jag $henne_1$ inte t_2 (bara hållit henne i handen)
kissed have I her not (just held her by the hand)

However, Holmberg (1997) argues that (i) does not actually involve remnant VP movement (because the object pronoun *henne* would have had to leave the VP without verb raising of *kysst* having taken place), but rather (exceptional) topicalization of a bare V^0 category. I will remain uncommitted as to whether such a drastic step is indeed necessary for (i), or whether the remnant movement analysis could be maintained after all – this second approach would minimally require a view of the dependence of object shift on verb fronting that is based not on a structural notion (c-command) but on a linear notion (precedence).

[12]Scrambling in Dutch is somewhat more restricted than its German counterpart; among other things, it normally may not change the D-structure order of arguments (cf. Neeleman (1994), among others). Interestingly, it turns out that Dutch incomplete category fronting is more restricted, too. For instance, examples like (1-bc) of the preface are not permitted in Dutch, in contrast to German (cf. Zwart (1993)). I will leave open the questions of how these two facts are to be accounted for, and whether the two phenomena could in fact be correlated, focussing on German throughout.

Constructions that are superficially similar to remnant topicalization in German can also be found in Celtic and Slavic languages, e.g., in Breton and Serbo-Croatian. A Breton example is given in (17) (cf. Borsley (1990, 82)):

(17) Lennet$_1$ en deus t_1 Tom al levr
read have$_3$.Sing/Masc Tom the book
'Tom has read the book.'

Under standard assumptions, VSO order in the Celtic languages is derived by raising of the finite verb out of the VP into a functional head position that precedes the subject position. Various kinds of categories can then be moved into a position in front of the finite verb, e.g., adverbs or objects. As shown in (17), what can also undergo fronting in this context is a participle, stranding its object in the VP. At first sight, this looks completely parallel to remnant topicalization of the German type, and this might constitute a problem given that Breton is a language that does not exhibit scrambling. However, as argued by Hendrick (1990, 128-131) and Schafer (1995, 12-13), there is reason to believe that (17) does not involve (remnant) VP movement, but rather movement of an X^0 category. I will confine myself to two pieces of evidence that immediately confirm this view: First, a construction that differs minimally from (17) in that it must involve fronting of a complete VP is in fact ungrammatical (see Borsley (1990, 82) and Hendrick (1990, 130)), which clearly suggests that participles can only be fronted via head movement in Breton; cf.:

(18) *[$_{VP}$ Al levr lennet]$_1$ en deus Tom t_1
the book read have$_3$.Sing/Masc Tom

And second, Borsley and Hendrick note that participle fronting of the type in (17) is always clause-bound in Breton. Again, this is a property of head movement, but not of XP movement (and indeed, remnant VP topicalization in German can apply long-distance, see subsection 3.3.3 of chapter 5 below).

Turning next to Serbo-Croatian, a relevant example that at first sight resembles remnant VP topicalization in German is given in (19) (see, e.g., Rivero (1991)):

(19) Čitao$_1$ sam t_1 knjigu
read have$_1$.Sing./Cl book
'I have read the book.'

Given that Serbo-Croatian has scrambling, a remnant movement analysis does not look impossible here. However, it seems that there is reason to believe that (19) does not have to involve a combination of scrambling and remnant topicalization, but can be analyzed in terms of X^0 movement that is phonologically triggered; cf. in particular Wilder & Čavar (1994, 61ff).

Summarizing, it seems that the remnant movement approach explains cross-linguistic variation with respect to the option of incomplete category fronting straightforwardly, which can thus be viewed as independent corroboration.

In the following three subsections, I will discuss an empirical consequence of the remnant movement approach: If there is an unbound trace in the fronted category that results from movement, the construction is expected to exhibit typical properties of movement. This prediction will be shown to be borne out by data involving extraction from NP (subsection 2.3), extraction from PP (subsection 2.4), and extraction from infinitives (subsection 2.5).

2.3. Remnant NPs

2.3.1. The Phenomenon

The following examples involve incomplete category topicalization with a nominal category (and not with a verbal category, as in the examples discussed up to now).

(20) a. [$_{\text{NP}}$ Ein Buch t$_1$]$_2$ hat Antje [$_{\text{PP}}$ über die Liebe]$_1$ t$_2$ gelesen
a book$_{acc}$ has Antje$_{nom}$ about the love read

b. [$_{\text{NP}}$ Ein Buch t$_1$]$_2$ hat [$_{\text{PP}}$ über die Liebe]$_1$ niemand t$_2$ gelesen
a book$_{acc}$ has about the love no-one$_{nom}$ read

In both (20-a) and (20-b), an NP-internal THEME argument (the PP *über die Liebe* ('about love')) fails to undergo fronting together with the noun it belongs to (*Buch* ('book')); this PP occurs after the subject in (20-a), and in front of the subject in (20-b). Under the remnant movement approach, an unbound scrambling trace of the THEME argument of the head noun is present in the fronted item in (20).

A similar situation appears in sentences like (21).

(21) [$_{\text{VP}}$ [$_{\text{NP}}$ Ein Buch t$_1$]$_2$ gelesen]$_3$ hat [$_{\text{PP}}$ darüber]$_1$ keiner t$_3$
a book$_{acc}$ read has about that no-one$_{nom}$

Here, a verb and its direct object undergo topicalization, leaving material behind, just as in (6) above. However, what is left behind in (21) is not another argument of the verb (as in (6), where the indirect object stays in the middle field), but rather the internal argument of the head noun of the direct object, as in (20-a) and (20-b). Again, the remnant movement approach predicts that a trace of this argument PP is present in the incomplete category that is topicalized. We are now led to expect that the construction in (20) and (21) will behave similarly to "regular" instances of extraction from NP in German; indeed, this seems to be the case, by and large. To illustrate this, I will first highlight some properties of extraction from NP in German, and argue that extraction from NP does indeed exist as a syntactic movement operation that leaves an NP-internal trace.

2.3.2. Extraction from NP

Consider some typical instances of extraction from NP in German, as in (22).

(22) a. Worüber$_1$ hat keiner [$_{NP}$ ein Buch t$_1$] gelesen ?
about what has no-one$_{nom}$ a book$_{acc}$ read
b. dass darüber$_1$ keiner je [$_{NP}$ ein Buch t$_1$] gelesen hat
that about that no-one$_{nom}$ ever a book$_{acc}$ read has

At least at first sight, it looks as though (22-a) involves *wh*-extraction of the PP *worüber* ('about what') from the object NP and as though (22-b) is a case of scrambling of the PP *darüber* ('about that') from the object NP. However, independently of the issue of remnant movement in (20) and (21), the literature disagrees about whether or not a trace t$_1$ is present in the examples in (22). Chomsky (1977) has proposed that English examples which completely parallel those in (22) only superficially give the impression of extraction from NP. Under this view, what looks like extraction from NP is actually the result of a reanalysis process that allows the PP to optionally be base-generated *outside* of the NP; see Fanselow (1987) (among others) for an account of examples like those in (22) in German along these lines. The purported reanalysis rule is then assumed to be subject to lexical variation, as shown in (23) (which is structurally identical to (22-a) – only the governing verb differs).

(23) *Worüber$_1$ hat keiner [$_{NP}$ ein Buch t$_1$] geklaut ?
about what has no-one$_{nom}$ a book$_{acc}$ stolen

In contrast to that view, Webelhuth (1987; 1992), Fanselow (1991), and Müller (1995) argue that extraction from NP in German has a number of properties that are problematic for a reanalysis approach and that suggest the existence of an NP-internal trace. I will briefly sketch some of these properties. First, extraction from NP never appears with subjects (of transitive or unergative verbs); cf. (24).

(24) *Worüber$_1$ hat [$_{NP}$ ein Buch t$_1$] Antje beeindruckt ?
about what has a book$_{nom}$ Antje$_{acc}$ impressed

Second, it never occurs with indirect objects (cf. (25-a)), even if the verb as such allows extraction from NP (cf. (25-b), where extraction from a direct object occurs).

(25) a. *[$_{PP}$ Worüber]$_1$ hat man [$_{NP}$ einem Buch t$_1$] einen Preis
about what has one$_{nom}$ a book$_{dat}$ an award$_{acc}$
gegeben ?
given
b. [$_{PP}$ Worüber]$_1$ hat man Antje [$_{NP}$ ein Buch t$_1$] gegeben ?
about what has one$_{nom}$ Antje$_{dat}$ a book$_{acc}$ given

Third, extraction from a definite NP exhibits a Specificity effect (see Chomsky (1981, 235), Mahajan (1992), and Webelhuth (1992, 170), among others); this is shown in (26).

(26) ?*[$_{PP}$ Worüber]$_1$ hat Antje [$_{NP}$ das Buch t$_1$] gelesen ?
about what has Antje$_{nom}$ the book$_{acc}$ read

Fourth, PP extraction is impossible if a pre-nominal possessor is present (a Specified Subject Condition (SSC) effect; see Chomsky (1973)):

(27) *[$_{PP}$ Worüber]$_1$ hat Antje [$_{NP}$ Karls Buch t_1] gelesen ?
about what has Antje$_{nom}$ Karl's$_{gen}$ book$_{acc}$ read

And finally, Freezing effects occur if the direct object undergoes scrambling, in the sense that the scrambled object becomes opaque for extraction; compare (28-a) to (28-b) (= (22-a)):

(28) a. *[$_{PP}$ Worüber]$_1$ hat [$_{NP}$ ein Buch t_1]$_2$ keiner t_2 gelesen ?
about what has a book$_{acc}$ no-one$_{nom}$ read
b. [$_{PP}$ Worüber]$_1$ hat keiner [$_{NP}$ ein Buch t_1] gelesen ?
about what has no-one$_{nom}$ a book$_{acc}$ read

All these island properties are indicative of movement, and given the presence of an NP-internal trace, they can be explained by invoking locality constraints (cf. chapter 2). On the other hand, under the reanalysis approach it is unclear how to account for the the data in (24) – (28).

Now, it is known that verbs like *lesen* ('read') in (28-b), in contrast to verbs like *klauen* ('steal') in (23), may occur in constructions in which the PP is present but the NP is either completely absent or realized only as a pronoun. This is generally taken to be the strongest argument in support of the base-generation approach to extraction from NP. Cf. (29-a) vs. (29-b).

(29) a. dass Fritz (?es) [$_{PP}$ über die Liebe] gelesen hat
that Fritz it about the love read has
b. *dass Fritz (es) [$_{PP}$ über die Liebe] geklaut hat
that Fritz it about the love stolen has

However, verbs like *geben* ('give') in German behave like *lesen* in that they permit extraction from NP (cf. (25-b)), but behave like *klauen* ('steal') in that they do not allow the NP to be pronominal or dropped (cf. (30)). Thus, the correlation breaks down, and with it goes the argument for reanalysis.

(30) *dass man (es) Antje [$_{PP}$ über die Liebe] gegeben hat
that one$_{nom}$ it$_{acc}$ Antje$_{dat}$ about the love given has

All in all, the reanalysis approach to extraction from NP turns out to be highly problematic from an *empirical* point of view. Given that there are also a number of substantial *conceptual* problems connected with this analysis (such as the formal nature of the reanalysis rule involved; cf. Webelhuth (1987), and also Baltin & Postal (1996, 135ff) for discussion), and given that the phenomenon of lexical variation with extraction from NP can be accounted for without (non-structure preserving) reanalysis processes (for instance, by invoking the structure-preserving notion of abstract incorporation; see Müller (1995, ch. 2)), I conclude that the examples in (22) involve genuine movement out of NP.

2.3.3. Remnant NPs

With this in mind, let us return to the examples in (20) and (21) that involve incomplete category topicalization; these examples are repeated here.

(31) a. [$_{\text{NP}}$ Ein Buch t$_1$]$_2$ hat Antje [$_{\text{PP}}$ über die Liebe]$_1$ t$_2$ gelesen
a book$_{acc}$ has Antje$_{nom}$ about the love read

b. [$_{\text{NP}}$ Ein Buch t$_1$]$_2$ hat [$_{\text{PP}}$ über die Liebe]$_1$ niemand t$_2$ gelesen
a book$_{acc}$ has about the love no-one$_{nom}$ read

c. [$_{\text{VP}}$ [$_{\text{NP}}$ Ein Buch t$_1$]$_2$ gelesen]$_3$ hat [$_{\text{PP}}$ darüber]$_1$ keiner t$_3$
a book$_{acc}$ read has about that no-one$_{nom}$

If an NP-internal trace t$_1$ is present in the examples in (31), the construction should exhibit the properties indicating movement that have just been discussed. This is indeed the case. Note first that (31-a) is ungrammatical if the governing verb *lesen* is replaced by *klauen* (cf. (23)):

(32) *[$_{\text{NP}}$ Ein Buch t$_1$]$_2$ hat Antje [$_{\text{PP}}$ über die Liebe]$_1$ t$_2$ geklaut
a book$_{acc}$ has Antje$_{nom}$ about the love stolen

Second, incomplete nominal categories that are subjects cannot be fronted (cf. (24)):

(33) a. *[$_{\text{NP}}$ Ein Buch t$_1$]$_2$ hat t$_2$ Antje [$_{\text{PP}}$ über die Liebe]$_1$ beeindruckt
a book$_{nom}$ has Antje$_{acc}$ about the love impressed

b. *[$_{\text{NP}}$ Ein Buch t$_1$]$_2$ hat t$_2$ [$_{\text{PP}}$ über die Liebe]$_1$ Antje beeindruckt
a book$_{nom}$ has about the love Antje$_{acc}$ impressed

Third, indirect object NPs cannot undergo topicalization if an internal argument of the head noun of the indirect object stays behind, even if the verb as such permits extraction from NP (with direct objects). This is just what one would expect under the remnant movement approach (cf. (25-a) and (25-b)):

(34) a. *[$_{\text{NP}}$ Einem Buch t$_1$]$_2$ hat man darüber$_1$ t$_2$ einen Preis gegeben
a book$_{dat}$ has one$_{nom}$ about that an award$_{acc}$ given

b. [$_{\text{NP}}$ Ein Buch t$_1$]$_2$ hat man Antje [$_{\text{PP}}$ über die Liebe]$_1$ t$_2$ gegeben
a book$_{acc}$ has one$_{nom}$ Antje$_{dat}$ about the love given

Fourth, a Specificity effect appears when an incomplete nominal category is fronted, again suggesting that this category contains a trace of movement (cf. (26)):

(35) ?*[$_{\text{NP}}$ Das Buch t$_1$]$_2$ hat Antje darüber$_1$ t$_2$ gelesen
the book$_{acc}$ has Antje$_{nom}$ about that read

And, finally, the construction is not immune to Specified Subject Condition (SSC) effects, which again is indicative of movement, i.e., of a trace in the fronted NP (cf. (27)):

(36) *[$_{NP}$ Karls Buch t_1]$_2$ hat Antje [$_{PP}$ über die Liebe]$_1$ t_2 gelesen
Karl's$_{gen}$ book$_{acc}$ has Antje$_{nom}$ about the love read

To sum up, the parallelism between cases of ordinary extraction from NP and cases involving fronting of an incomplete nominal category is explained directly under the remnant movement approach; remnant NP topicalization behaves just as we would expect it to.[13]

2.4. Remnant PPs

PPs in German pattern with NPs rather than with VPs in that extraction from PP is possible only under certain circumstances. Hence, it does not come as a surprise that more or less the same conclusions can be drawn as in the preceding subsection.

As is well known (see, e.g., van Riemsdijk (1978)), extraction of an NP from a PP (i.e., P-stranding) in languages like Dutch and German is possible only if the NP is an R-pronoun (like *wo* ([+wh]) and *da* ([–wh]) in German) and not a full lexical NP. This is shown for German in (37-a) vs. (37-b):[14]

(37) a. Wo$_1$ hat keiner [$_{PP}$ t_1 mit] gerechnet ?
where has no-one$_{nom}$ with counted

b. *[$_{NP}$ Dieser Sache]$_1$ hat keiner [$_{PP}$ mit t_1] gerechnet
this thing$_{dat}$ has no-one$_{nom}$ with counted

At the moment, it is not important exactly how we account for the prohibition against extraction of full lexical NPs from PP in German (and Dutch) nor why PPs are not islands for extraction of R-pronouns in these languages. The only important thing in this context is that there is a condition on movement which is violated in (37-b) but met in (37-a) (following Bennis & Hoekstra (1984), Koster (1987), and Grewendorf (1989), one might assume that t_1 is selected in the "wrong direction" in (37-b), but see the last footnote). With this in mind, consider the instances of incomplete category fronting in (38):

---

[13]It should be noted, though, that the construction at hand does not seem to exhibit a Freezing effect, in striking contrast to ordinary instances of extraction from NP like (28-a). I will come back to this in section 3, and then discuss the issue at length in chapter 4.

[14]Along with most of the literature on this topic, I assume here that extraction of an R-pronoun from a PP in German is an instance of *postposition* stranding, not *preposition* stranding, but see Müller (1995, 60-64 & 394:note 44) for arguments against this view. However, nothing hinges on this in the present context.

(38) a. ?[$_{VP}$ [$_{PP}$ t$_1$ Mit] gerechnet]$_2$ hat da$_1$ keiner t$_2$
with counted has there no-one$_{nom}$
b. *[$_{VP}$ [$_{PP}$ t$_1$ Mit] gerechnet]$_2$ hat [$_{NP}$ dieser Sache]$_1$ keiner t$_2$
with counted has this thing$_{dat}$ no-one$_{nom}$

In (38-a), the postposition *mit* ('with') undergoes topicalization together with the participle, leaving the R-pronoun *da* in the middle field, in front of the subject. The resulting sentence is well formed, or at least much better than its ungrammatical counterpart in (38-b), which differs minimally from (38-a) in that the R-pronoun *da* is replaced by the full lexical NP *dieser Sache* ('this thing').[15] Thus, the sentences in (38) completely parallel those in (37). This state of affairs strongly suggests that the fronted VP in (38-a) contains a trace of the R-pronoun, just as (37-a) does. In conclusion, the very existence of constructions like the one in (38-a) again supports the remnant movement approach to incomplete category fronting.

So far, so good. It should be noted, however, that with extraction from PP there is an intervening factor that blurs the overall picture. In contrast to what was the case with NPs (cf. (31-a) and (31-b)), it does not appear to be possible to front just the PP, as in (39) (cf. den Besten & Webelhuth (1990, 82)):

(39) *[$_{PP}$ t$_1$ Mit]$_2$ hat da$_1$ keiner t$_2$ gerechnet
with has there no-one$_{nom}$ counted

Still, there is reason to believe that the illformedness of (39) does not have anything to do with the issue of remnant movement. To see this, consider (40):

(40) a. *[$_{PP}$ Mit]$_1$ hat gestern keiner t$_1$ gespielt
with has yesterday no-one$_{nom}$ played
b. Gestern spielte keiner [$_{PP}$ mit]
yesterday played no-one with

Assuming along with Emonds (1976), Kayne (1985), and others that separable prefixes like *mit* in *mitspielen* (cf. (40-b)) are base-generated as intransitive PPs, I see no a

---

[15]Note that (38-a) is not judged grammatical in den Besten & Webelhuth (1990, 86). However, examples like (38-a) or (i-a)

(i) a. ?[$_{VP}$ [$_{PP}$ t$_1$ Für] gekonnt]$_2$ hat er da$_1$ wirklich nichts t$_2$
for could$_{Part}$ has he there really nothing
'He really was not responsible for that.'
b. *[$_{VP}$ [$_{PP}$ Für t$_1$] gekonnt]$_2$ hat er [$_{NP}$ diese Entwicklung]$_1$ wirklich nichts t$_2$
for could$_{Part}$ has he this development really nothing
'He really was not responsible for this development.'

are at worst slightly marginal for many speakers (in particular, for those belonging to northern varieties of German), and occur regularly in colloquial speech. In any event, the contrast to examples like (38-b) or (i-b) should be evident for all speakers of German that permit postposition stranding at all.

priori reason why they should not be able to undergo topicalization, as in (40-a). But then, whatever eventually accounts for why P fronting is impossible in (40-a) is likely to also rule out incomplete category fronting in (39), so that we can conclude that the evidence to be gained from (39) does not bear on the issue of remnantness.[16] That said, let me now turn to the third set of extraction data.

2.5. Coherent Infinitives

Whereas scrambling from finite clauses is strictly prohibited in German (cf. Ross (1967), Fanselow (1990), and Müller & Sternefeld (1993), among many others), the case is different with non-finite clauses. Certain kinds of infinitival clauses in German are transparent for scrambling (so-called "coherent" infinitives), whereas others are not ("incoherent" infinitives); cf. Bech (1955/57). The contrast between coherent and incoherent infinitives is shown in (41).

(41) a. dass [$_{\text{NP}}$ das Buch]$_1$ keiner [$_\alpha$ t$_1$ zu lesen] versucht hat
that the book$_{acc}$ no-one$_{nom}$ to read tried has

b. ?*dass [$_{\text{NP}}$ das Buch]$_1$ keiner [$_\alpha$ t$_1$ zu lesen] abgelehnt hat
that the book$_{acc}$ no-one$_{nom}$ to read refused has

Evidently, the only difference between (41-a) and (41-b) concerns the choice of the matrix verb – a verb like *versuchen* ('try') in German makes the infinitive transparent and thereby permits scrambling out of the embedded clause, whereas a verb like *ablehnen* ('refuse') does not create transparency. This systematic lexical variation in German has been accounted for in various ways (recent work includes Grewendorf (1987), Stechow & Sternefeld (1988), Sternefeld (1990), Fanselow (1991), Haider (1991; 1993), Sabel (1995), and Müller & Sternefeld (1995)). However, since a commitment to one of the competing theories of coherent infinitives in languages like German is – for the most part – not necessary for my present purposes, I will not take a firm stand on this issue here. Similarly, for the time being I will not commit myself with regard to the categorial status of coherent and incoherent infinitives in German (i.e., whether they are to be analyzed as CPs, IPs, VPs, or projections of some other functional head); the categorial status certainly bears on the issue of transparency. The only relevant thing in the present context is that scrambling from an infinitival clause

---

[16] Note also that the hypothesis that the illformedness of (39) is not due to a general constraint against remnant PP movement is reinforced by the fact that analogous constructions are possible in Italian; consider (i) (from Cinque (1990, 176)):

(i) [$_{\text{PP}}$ Addosso t$_1$]$_2$ non gli$_1$ è caduta t$_2$
on not him$_{dat}$ (she) is fallen

The only relevant difference to the German example in (39) apparently is that the antecedent of the unbound trace has not undergone scrambling, but rather cliticization. I will return to this kind of construction in chapter 5 (section 5).

is subject to lexical variation; it is possible with certain embedding verbs, but not with others.

Interestingly now, lexical variation in extraction options as it appears in (41) is mirrored in cases of incomplete category fronting (cf. Fanselow (1989, 4), Grewendorf & Sabel (1994, 302) and Müller & Sternefeld (1995, 62)). Relevant examples are given in (42):

(42) a. [$_\alpha$ t_1 Zu lesen]$_3$ hat [$_{NP}$ das Buch]$_1$ keiner t_3 versucht
to read has the book no-one tried

b. *[$_\alpha$ t_1 Zu lesen]$_3$ hat [$_{NP}$ das Buch]$_1$ keiner t_3 abgelehnt
to read has the book no-one refused

In both sentences, the object NP of the embedded infinitive shows up in front of the subject, in the matrix clause, and the infinitive itself has undergone topicalization. The result is well formed with the verb *versuchen* ('try') in (42-a), which permits a coherent construction, and ill formed with the verb *ablehnen* ('refuse') in (42-b), which is neccessarily construed incoherently. Thus, the illformedness of (42-b) and the wellformedness of (42-a) correspond to expectations, given the remnant movement analysis.

Next, consider the examples in (43), which differ minimally from those in (42) in that the object NP of the infinitive shows up in a position following the matrix subject:

(43) a. [$_\alpha$ t_1 Zu lesen]$_3$ hat keiner [$_{NP}$ das Buch]$_1$ t_3 versucht
to read has no-one the book tried

b. *[$_\alpha$ t_1 Zu lesen]$_3$ hat keiner [$_{NP}$ das Buch]$_1$ t_3 abgelehnt
to read has no-one the book refused

Again, the result is well formed with the embedding verb *versuchen* in (43-a), and ill formed with *ablehnen* in (43-b); and, again, it seems (at first sight at least) that this does not come as a surprise under the remnant movement approach.[17]

---

[17]There is, however, a complication, pointed out by Fanselow (1993, 28). In order to guarantee the illformedness of (43-b), one must ensure that the NP *das Buch* is indeed scrambled out of the embedded infinitival clause, and not simply scrambled within it. To see this, suppose that NP_1 in (43-b) is not actually moved out of α, but scrambled locally to α, creating a two-segment category, and the VP headed by the verb *zu lesen* ('to read') is then topicalized. This is shown in (i):

(i) *[$_{VP}$ t_1 Zu lesen]$_3$ hat keiner [$_\alpha$ [$_{NP}$ das Buch]$_1$ t_3] abgelehnt
to read has no-one the book refused

If NP_1 is scrambled within the incoherent infinitive, the fact that the incoherent infinitive α does not permit scrambling out of it is irrelevant. One might hope, then, that topicalization of VP in (i) is still blocked by the presence of α. However, it is a fact that incoherent infinitives, although they block scrambling into the matrix clause, are not an obstacle for other movement operations, like topicalization and *wh*-movement. This is shown in (ii):

Consider next some cases of topicalization of an incomplete verbal category where the fronted item contains an infinitive of the type in (43):

(44) a. [$_{\text{VP}}$ [$_\alpha$ t$_1$ Zu lesen]$_3$ versucht]$_2$ hat keiner [$_{\text{NP}}$ das Buch]$_1$ t$_2$
to read tried has no-one the book$_{acc}$

b. *[$_{\text{VP}}$ [$_\alpha$ t$_1$ Zu lesen]$_3$ abgelehnt]$_2$ hat keiner [$_{\text{NP}}$ das Buch]$_1$ t$_2$
to read refused has no-one the book$_{acc}$

Again, the same contrast shows up as between (42-a) and (42-b), or between (43-a) and (43-b). And again, this contrast follows under the remnant movement approach. Unlike what was the case with string-vacuous NP scrambling in a position following the subject in (43-b), no problem appears this time (cf. the previous footnote): NP$_1$ in (44-b) must be scrambled out of the incoherent infinitive α, thereby violating a locality constraint, because α-internal scrambling invariably would have meant that NP$_1$ undergoes fronting together with the two verbs *zu lesen* ('to read') and *abgelehnt* ('refused').

Finally, note that the coherence/incoherence distinction obtains with even more complex examples, as in (45):

(45) a. [$_{\text{VP}}$ [$_\alpha$ Dem Peter t$_1$ zu geben] versucht]$_2$ hat [$_{\text{NP}}$ das Buch]$_1$
ART Peter$_{dat}$ to give tried has the book$_{acc}$
keiner t$_2$
no-one$_{nom}$

b. *[$_{\text{VP}}$ [$_\alpha$ Dem Peter t$_1$ zu geben] abgelehnt]$_2$ hat [$_{\text{NP}}$ das Buch]$_1$
ART Peter$_{dat}$ to give refused has the book$_{acc}$
keiner t$_2$
no-one$_{nom}$

(ii) a. [$_{\text{NP}}$ Das Buch]$_1$ hat keiner [$_\alpha$ t$_1$ zu lesen]$_3$ abgelehnt
the book$_{acc}$ has no-one$_{nom}$ to read refused

b. Was$_1$ hat keiner [$_\alpha$ t$_1$ zu lesen]$_3$ abgelehnt ?
what$_{acc}$ has no-one$_{nom}$ to read refused

For present purposes, it is immaterial why exactly incoherent infinitives do not block specifier movements like topicalization and *wh*-movement (cf. chapter 6); the only important thing in this context is that given this observation, (i) very much looks like a legitimate derivation and seems to be predicted to be well formed, contrary to fact. Thus, there must be some independent factor that guarantees that scrambling of NP$_1$ in (43-b) (and in (43-a), for that matter) must end up in a position external to the embedded non-finite clause, and cannot take place within the incoherent infinitive; this is indeed the conclusion arrived at in Grewendorf & Sabel (1994, 290-291), Müller & Sternefeld (1995, 65-73), and Zwart (1993, 108). I will not go into these analyses here, mainly because that would require a commitment to an articulated theory of the structure of coherent and incoherent infinitives. (Incidentally, the first two analyses build on Baker's (1988) approach that crucially posits VP raising to an infinitive-internal specifier position.) It may suffice to bear in mind that (43-b) cannot be analyzed in terms of local scrambling of NP$_1$ under these analyses, and that, therefore, the only possibility that remains is long-distance scrambling of NP$_1$ into the matrix clause, an option that is not available here (for lexical reasons).

In (45), a VP comprising (a) the embedded infinitival verb, (b) its indirect object, and (c) the embedding participle is topicalized; the direct object NP of the embedded verb is scrambled to a position in front of the matrix subject. Such topicalization is legitimate if the fronted participle permits a coherent construction, as in (45-a), and illegitimate if the fronted participle does not, as in (45-b). Again, this follows under the remnant movement approach.

Summarizing the last three subsections, there are typical restrictions on extraction from NP, PP, and infinitivals, and these restrictions can be taken to be indicative of movement, i.e., of the presence of traces. The fact that basically all of these restrictions (with the exception of Freezing effects, see note 13) are preserved under incomplete category fronting is just what we would expect under the remnant movement approach, and can therefore be viewed as corroborating it.

3. Peculiar Properties

In the last section, I have shown that there is evidence suggesting that something like remnant movement does exist – fronted "incomplete categories" behave as though they contain (unbound) traces in many respects. However, closer scrutiny reveals that remnant movement has a number of striking properties that set it apart (at first sight at least) from other, well-established cases of movement. In what follows, I would like to highlight four such properties of remnant movement, concerning (a) the Proper Binding Condition, (b) the absence of Freezing effects, (c) unexpected asymmetries between types of movement, and (d) an otherwise undocumented restriction on the kind of trace that may show up in remnant categories.

3.1. Proper Binding

Maybe the most obvious peculiarity of remnant movement constructions is that they violate the Proper Binding Condition (PBC). Various versions of the PBC have been proposed (see, among others, Fiengo (1977), May (1977), Barss (1986), Saito (1989), Lasnik & Saito (1992), and chapter 3 below). A simple version looks as follows.

(46) *The Proper Binding Condition* (PBC):
Traces must be bound.

Among other things, the PBC precludes syntactic lowering, as in (47) (assuming for the time being that the PBC applies at S-structure and at LF).

(47) *John asked t_1 [$_{CP}$ who_1 Mary saw Bill]

The matrix verb *ask* selects a *wh*-clause, and this selection requirement is met in (47) by S-structure lowering of the matrix object *who* into the embedded SpecC position. Such cases of downward *wh*-movement in the syntax are correctly excluded by the PBC.

Similarly, the PBC makes some non-trivial empirical predictions for topicalization options. Consider, e.g., an example like (48) in German.[18]

(48) *dass Fritz t_1 sagte [$_{CP}$ dem Peter$_1$ habe [$_{IP}$ keiner die Claudia
that Fritz said ART Peter$_{dat}$ has$_{subj}$ no-one$_{nom}$ ART Claudia$_{acc}$
gesehen]]
seen

A requirement that holds in V/2 languages like German is that a V/2 head needs a filled specifier (which can be a lexical category or a trace). This requirement is fulfilled in (48) by downward topicalization of the matrix object *dem Peter* into the embedded SpecV/2 position. To rule out such lowering, the PBC can successfully be invoked – from this condition, it follows that SpecV/2 in (48) must be occupied either by some XP that originates in the embedded CP (as in (49-a)), or by an expletive pronoun *es* (as in (49-b)):

(49) a. dass Fritz dem Peter$_1$ sagte [$_{CP}$ die Claudia$_2$ habe [$_{IP}$ keiner
that Fritz ART Peter$_{dat}$ said ART Claudia$_{acc}$ has$_{subj}$ no-one$_{nom}$
t_2 gesehen]]
seen

b. dass Fritz dem Peter$_1$ sagte [$_{CP}$ es habe [$_{IP}$ keiner die
that Fritz ART Peter$_{dat}$ said EXPL has$_{subj}$ no-one$_{nom}$ ART
Claudia gesehen]]
Claudia$_{acc}$ seen

Turning now to a typical remnant movement construction such as (2) (repeated here as (50)), it is obvious that given the PBC in (46), we would expect ungrammaticality to arise (just as it does in (47) and (48)), contrary to fact: (50) is grammatical even though t_2 is not bound by NP_2.

(50) [$_{VP}$ t_2 Gelesen]$_1$ hat [$_{IP}$ [$_{NP}$ das Buch]$_2$ [$_{IP}$ keiner t_1]]
read has the book$_{acc}$ no-one$_{nom}$

Thus, the fact that remnant movement creates traces which are unbound at S-structure and may thereby systematically violate the PBC calls for an explanation.

3.2. Freezing

In general, extraction of one XP α from another XP β is possible only if β occurs in situ, as is depicted schematically in (51-a), and has not undergone movement itself, as shown in (51-b).

---

[18] I assume here, in line with the standard view (see, e.g., den Besten (1989), Vikner (1995)), that topicalization in German is movement to SpecC and that V/2 movement ends up in C. See, however, chapter 2 for a different view.

(51) a. ... α_1 ... [$_\beta$ ... t_1 ...]$_2$
b. *... α_1 ... [$_\beta$ ... t_1 ...]$_2$ ... t_2

More generally, the following generalization seems to be descriptively adequate, by and large (cf. Ross (1967/1986, 173) and Wexler & Culicover (1980, 119)):

(52) *Freezing*:
At S-structure, a trace t may not be included in a moved XP (i.e., an XP that binds a trace) if the antecedent of t is excluded by XP.

Henceforth, I will refer to this as the Freezing generalization, leaving the question of how it can be derived from independent principles open for the time being (see chapter 4). In the last section, we already encountered one instance of Freezing – an object NP that is transparent for extraction in situ invariably blocks extraction in a derived scrambling position; recall the examples in (28), repeated here:

(53) a. *[$_{\text{PP}}$ Worüber]$_1$ hat [$_{\text{NP}}$ ein Buch t_1]$_2$ keiner t_2 gelesen ?
about what has a book$_{acc}$ no-one$_{nom}$ read
b. [$_{\text{PP}}$ Worüber]$_1$ hat keiner [$_{\text{NP}}$ ein Buch t_1] gelesen ?
about what has no-one$_{nom}$ a book$_{acc}$ read

Similarly, whereas a VP in situ is not an island for extraction out of it, it turns into an island if it is moved. This is shown for topicalization in German in (54). Thus, (54-a) indicates that topicalization of a complete VP in German is possible; (54-b) indicates that a VP is transparent for extraction in situ; finally, (54-c) exhibits the Freezing effect with VP topicalization – there is no extraction from a topicalized VP.

(54) a. Ich denke [$_{\text{CP}}$ [$_{\text{VP}}$ das Buch gelesen]$_2$ hat keiner t_2]
I think the book read has no-one
b. [$_{\text{NP}}$ Was]$_1$ denkst du [$_{\text{CP}}$ t'_1 hat keiner [$_{\text{VP}}$ t_1 gelesen]$_2$] ?
what think you has no-one read
c. *[$_{\text{NP}}$ Was]$_1$ denkst du [$_{\text{CP}}$ [$_{\text{VP}}$ t_1 gelesen]$_2$ hat keiner t_2] ?
what think you gelesen has no-one

The illformedness of examples like (53-a) and (54-c) corresponds exactly to what one would expect, given the Freezing generalization in (52); in these cases, a trace t_1 shows up in the fronted items (a scrambled NP and a topicalized VP, respectively), and the antecedent of this trace (PP_1/NP_1) is excluded by the moved XP.

However, typical remnant movement constructions like those in (55) (= (50) and (31-a)) contradict the otherwise well-established Freezing generalization.

(55) a. [$_{\text{VP}}$ t_2 Gelesen]$_1$ hat [$_{\text{IP}}$ [$_{\text{NP}}$ das Buch]$_2$ [$_{\text{IP}}$ keiner t_1]]
read has the book$_{acc}$ no-one$_{nom}$
b. [$_{\text{NP}}$ Ein Buch t_1]$_2$ hat Antje [$_{\text{PP}}$ über die Liebe]$_1$ t_2 gelesen
a book$_{acc}$ has Antje$_{nom}$ about the love read

Here, an unbound trace (t_2 in (55-a), t_1 in (55-b)) shows up in a moved item (a topicalized VP and a topicalized NP, respectively), and the antecedent of this unbound trace is excluded by the moved item in both cases. Still, a Freezing effect fails to appear. Thus, as with the PBC, remnant movement behaves differently from other, standard instances of movement, in that it demands a systematic exception to the Freezing generalization (52). Hence, whatever eventually derives the effects of the Freezing generalization must not simultaneously rule out remnant movement as in (55).

3.3. Movement Type Asymmetries

A third peculiar property of remnant movement is that asymmetries between movement types seem to arise that do not show up with regular instances of movement. Thus, whereas the movement type topicalization seems to be able to affect remnant XPs (in fact, all the examples of remnant movement discussed so far are instances of remnant topicalization), the movement type scrambling generally may not affect remnant XPs (cf. Fanselow (1991, 133; 1992), Grewendorf & Sabel (1994, 284ff), Frank, Lee & Rambow (1992, 152), Stechow (1992, 222ff), and Haider (1993, 234ff)). To see this, recall from subsection 2.5 that topicalization of a coherent (i.e., remnant) infinitive is permitted in German (cf. (42-a), repeated here as (56)):

(56) [$_\alpha$ t_1 Zu lesen]$_3$ hat [$_{NP}$ das Buch]$_1$ keiner t_3 versucht
to read has the book$_{acc}$ no-one tried

In striking contrast to this, scrambling of a coherent infinitive is strictly ruled out; cf.:

(57) *dass [$_\alpha$ t_1 zu lesen]$_3$ [$_{NP}$ das Buch]$_1$ keiner t_3 versucht hat
that to read the book$_{acc}$ no-one tried has

As shown in (58-a) (with topicalization) and (58-b) (with scrambling), such an asymmetry does not arise in the case of movement of complete infinitival clauses that do not contain unbound traces:[19]

(58) a. [$_\alpha$ Das Buch zu lesen]$_3$ hat keiner t_3 versucht
the book$_{acc}$ to read has no-one tried
b. dass [$_\alpha$ das Buch zu lesen]$_3$ keiner t_3 versucht hat
that the book$_{acc}$ to read no-one tried has

Other movement types pattern either with topicalization, in that they permit remnant movement, or with scrambling, in that they do not.

[19] I think that the contrast between (57) and (58-b) eventually provides an additional argument in favour of the remnant movement approach to incomplete category fronting. The reason is that it is highly unclear how to derive the illformedness of (57) if α is the same kind of constituent in (57) as in (58-b), i.e., does not contain a trace in either of the examples.

Thus, *wh*-movement can successfully apply to remnant XPs, just as it can apply to full XPs. This is shown for remnant NP *wh*-movement in German in (59-a) (compare *wh*-movement of a complete NP in (59-b)):

(59) a. [$_{\text{NP}}$ Was für ein Buch t$_1$]$_2$ hast du [$_{\text{PP}}$ über die Liebe]$_1$ t$_2$ gelesen ?
what for a book$_{acc}$ have you about the love read

b. [$_{\text{NP}}$ Was für ein Buch über die Liebe]$_2$ hast du t$_2$ gelesen ?
what for a book$_{acc}$ about the love have you read

On the other hand, it appears that remnant XPs cannot undergo left dislocation in German, in contrast to complete XPs (cf. Truckenbrodt (1992)). This is shown in (60-a) (with remnant VP left dislocation) vs. (60-b) (with left dislocation of a complete VP).

(60) a. *[$_{\text{VP}}$ t$_1$ Gelesen], das hat das Buch$_1$ keiner t
read that has the book no-one

b. [$_{\text{VP}}$ Das Buch gelesen], das hat keiner t
the book read that has no-one

Thus, it appears that certain movement types (like topicalization and *wh*-movement) permit remnant movement, whereas others (like scrambling and left dislocation) do not. These movement type asymmetries in the case of remnant XPs are a priori unexpected; as in the cases of the PBC and the Freezing generalization, remnant movement behaves differently from regular instances of movement, and this fact needs to be explained.

3.4. Unbound Intermediate Traces

Finally, den Besten & Webelhuth (1990, 86), Fanselow (1993, 10), Grewendorf (1994, 31ff), and Bayer (1996, 242) have observed that remnant movement differs from other kinds of movement in that it imposes restrictions on the kinds of traces contained in the moved category. One particularly interesting restriction is this: Whereas standard instances of movement may of course give rise to configurations containing intermediate traces, intermediate traces that become unbound as a result of remnant movement appear to be illicit. To see this, consider first (61), which exemplifies a typical instance of successive-cyclic extraction in German:

(61) (Ich weiß nicht) wen$_1$ sie gesagt hat [$_{\text{CP}}$ t$'_1$ dass Fritz t$_1$ liebt]
I know not who$_{acc}$ she$_{nom}$ said has that Fritz$_{nom}$ loves

For reasons of locality, an intermediate trace t$'_1$ is created in the embedded SpecC position, and evidently, this trace does not pose any problem. Next, (62) shows that a finite CP argument that does not contain any unbound traces may undergo long-distance topicalization across a *wh*-island in German, with only weak (Subjacency-like) ungrammaticality arising (see Fanselow (1987)):[20]

(62) ??[$_{CP}$ Daß Fritz Caroline liebt]$_2$ weiß ich nicht [$_{CP}$ ob er t_2
that Fritz$_{nom}$ Caroline$_{acc}$ loves know I$_{nom}$ not whether he$_{nom}$
zugeben würde]
admit would

However, such long-distance CP topicalization across a *wh*-island becomes completely impossible if the CP contains an unbound trace, as in (63):

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

It is not a priori clear why (63) is much worse than (62), i.e., much worse than we would expect if only a weak Subjacency violation were involved. As shown in (64-a), the illformedness of (63) cannot possibly be due to a general prohibition against remnant movement across a *wh*-island (created by the antecedent of the unbound trace) – in (64-a), a remnant VP undergoes long-distance topicalization across a *wh*-island, and the result is no worse than the analogous case with topicalization of a complete VP (as in (64-b)):

(64) a. ??[$_{VP}$ t_1 Auf den Mund geküßt]$_2$ weiß ich nicht [$_{CP}$ wen$_1$ sie t_2
on the mouth kissed know I$_{nom}$ not who$_{acc}$ she$_{nom}$
hat
has

b. ??[$_{VP}$ Den Fritz auf den Mund geküßt]$_2$ weiß ich nicht [$_{CP}$ ob
ART Fritz$_{acc}$ on the mouth kissed know I$_{nom}$ not whether
sie t_2 hat
she$_{nom}$ has

The crucial difference between (63) and (64-a) seems to be that the remnant CP in (63) contains an intermediate unbound trace t'_1 in the SpecC position, whereas the remnant VP in (64-a) only contains the initial trace in object position. Thus, it appears that movement may not apply to a remnant XP if, as a result of this movement, this XP contains an unbound intermediate trace.[21]

[20]I assume here that finite CP arguments are base-generated to the left of the verb in German and then (normally) undergo extraposition. In the present context, this assumption is not important; see chapter 4 below for extensive discussion.

[21]Note that if an intermediate trace t'_1 were not present in (63), strong ungrammaticality would be expected regardless of the issue of remnant movement, due to a strong violation of locality requirements; see Müller & Sternefeld (1993).

This hypothesis is further corroborated by the illformedness of examples like (65-a) (compare (65-b)), in which a VP has undergone long-distance topicalization across a *wh*-island; here, the VP contains a CP from which successive-cyclic *wh*-movement has applied (see Fanselow (1993, 10)).

(65) a. *[$_{\text{VP}}$ Gesagt [$_{\text{CP}}$ t_1' dass Fritz t_1 liebt]]$_3$ weiß ich nicht [$_{\text{CP}}$ wen$_1$
said that Fritz$_{nom}$ loves know I$_{nom}$ not who$_{acc}$
er t_3 hat]
he$_{nom}$ has

b. ??[$_{\text{VP}}$ Gesagt [$_{\text{CP}}$ dass Fritz Caroline liebt]]$_3$ weiß ich nicht [$_{\text{CP}}$
said that Fritz$_{nom}$ Caroline$_{acc}$ loves know I$_{nom}$ not
ob er t_3 hat]
whether he$_{nom}$ has

In conclusion, it appears that there is a prohibition against unbound intermediate traces created by remnant movement which is responsible for the ungrammaticality of the sentences in (63) and (65-a). However, to the extent that this generalization is correct, it is an open question how it can be derived.

4. Conclusion

In section 2, I have given a number of (conceptual and empirical) arguments in support of the idea that incomplete category fronting in German involves remnant movement, an operation that creates an unbound trace in the moved item:

1. The postulation of remnant movement seems inevitable according to well-established restrictions on movement and projection (subsection 2.1).
2. The remnant movement hypothesis offers a straightforward account of cross-linguistic variation (subsection 2.2).
3. Constructions involving fronting of incomplete NPs (and VPs containing incomplete NPs) mirror the properties of constructions involving extraction from NP (subsection 2.3).
4. Constructions involving fronting of VPs containing incomplete PPs mirror the properties of constructions involving extraction from PP (subsection 2.4).
5. Constructions involving fronting of incomplete infinitives (and VPs containing incomplete infinitives) mirror the properties of constructions involving scrambling from infinitives (subsection 2.5).

Ideally, one would hope that nothing more has to be said, i.e., that remnant movement behaves in all respects like regular, standard instances of movement. However, as shown in section 3, this is not the case. Remnant movement exhibits a number of peculiar properties that at first sight set it apart from ordinary extraction, and that are unexpected within a standard theory of movement. Most notably, remnant movement constructions raise the following questions:

1. Why can remnant movement violate the Proper Binding Condition? (subsection 3.1)
2. Why is remnant movement an exception to the Freezing generalization? (subsection 3.2)
3. Why does remnant movement show movement type asymmetries? (subsection 3.3)
4. Why can remnant movement not create unbound intermediate traces? (subsection 3.4)

The main goal in the remainder of this monograph will be to provide answers to these questions that do not rest on construction-specific assumptions. In particular, I want to show that a strictly derivational approach to movement along the lines of the minimalist program that has developed out of the Principles and Parameters approach (cf. Chomsky (1991; 1993; 1995)) can successfully account for the a priori surprising properties of remnant movement.[22] Thus, before I tackle the problems posed by the four questions (in chapters 3 – 6), I will leave the issue of remnant movement aside for a while and turn to a derivational theory of movement based on minimalist assumptions in the following chapter; this theory will then be the basis for analyzing remnant movement.

[22]This does not mean that every specific proposal of Chomsky's minimalist program can be adopted. For instance, in a context that is unrelated to remnant movement constructions of the type I am interested in here, Chomsky (1995, 365) introduces a constraint that strictly blocks movement of any category that contains a trace (his ex. (200)). As observed by Fanselow (1997), adherence to this constraint implies giving up the remnant movement approach to incomplete category fronting. However, independently of the issue of remnant movement, it seems that such a constraint is untenable – among other things, it incorrectly rules out topicalization of a complex NP that dominates a clause in which some other movement operation (topicalization, V/2, etc.) has applied, or topicalization of a clause in which *wh*-movement has taken place.

Chapter 2

Movement Theory

1. Introduction

Movement in the minimalist program (Chomsky (1991; 1993; 1995); Chomsky & Lasnik (1993)) can be viewed as being constrained by various interacting factors.[1] First, movement theory comprises *locality* constraints. These constraints ensure that movement operations of any kind must be sufficiently "local." Second, minimalist movement theory crucially makes use of *economy* constraints; these constraints require that movement operations be "economical," in a sense that will be clarified in what follows. And third, I will assume that there is a third type of constraint, viz., *cyclicity* conditions. These constraints regulate the interaction of different movement operations.

One typical locality constraint is the prohibition against crossing a barrier (see Chomsky (1986), Chomsky & Lasnik (1993)); another is Rizzi's (1990) Relativized Minimality, which shows up in a slightly different form in the minimalist program as Minimize Chain Links; cf. Chomsky (1993; 1995), Chomsky & Lasnik (1993)). Relativized Minimality turns out not to be immediately relevant to the issue of remnant movement, and therefore I will not explicitly discuss it in what follows, leaving open the question of whether its effects can be fully derived by other constraints, or whether

---

[1] Note that what follows is one specific version of minimalist movement theory. Moreover, this version will turn out to be fairly conservative in several respects and to differ from the "standard" approach(es) in others (particularly in sections 4 – 6 below). Two general differences are particularly obvious but, I think, not crucial for what will be said about remnant movement in chapters 3–6. First, I do not adopt the assumption (put forward in Chomsky (1993; 1995)) that structure-building (Merge) and movement (Move) operations may systematically alternate during a derivation (cf. subsection 4.1.5 below), but rather adhere to the more traditional view according to which movement can take place only after all structure-building operations have been completed (with the standard exception of those instances of movement that involve adjunction, which is structure-building by itself). And second, I do not take over Chomsky's (1995) suggestion that movement operations actually apply to features (and not to categories, as was standardly assumed), with pied piping of other material giving the impression of X^0 or XP movement.

it is indeed required as a genuine principle of grammar (all claims of this study are in principle compatible with the existence of such a constraint).[2]

Economy constraints that have been proposed in the minimalist program are the Fewest Steps condition (cf. Chomsky (1991; 1993; 1995), Epstein (1992), Collins (1994), Fox (1994), and Kitahara (1995), among others); Last Resort (cf. Chomsky (1993; 1995), Collins (1997)); Procrastinate (cf. Chomsky (1993)); and the Shortest Path Condition (cf. Chomsky (1993; 1995), Kitahara (1993), and Collins (1994)).

Finally, as far as cyclicity is concerned, the relevant constraints which will be discussed are the Strict Cycle Condition (which was introduced in Chomsky (1973)) and a prohibition against improper movement (cf. Chomsky (1981), Fukui (1993a), Müller & Sternefeld (1993; 1996), and Takahashi (1994)).

In what follows (sections 2–4), I will introduce locality, economy, and cyclicity constraints in turn, focusing on their interaction; in doing so, I will concentrate on constructions involving A-bar movement (*wh*-movement, topicalization, scrambling). After that, I will suggest several extensions of the system that has emerged by then (sections 5 & 6).

2. Locality Constraints: The Barriers Condition

The crucial locality constraint to be discussed here is the prohibition against crossing a barrier (cf. Chomsky (1986)). This prohibition used to be part of the ECP and the Subjacency condition, but since these two conditions are absent in the minimalist framework, it seems as though the prohibition must be stated as such (cf. Chomky & Lasnik (1993, 540ff) and Chomsky (1995, 264)). It can be formulated as in (1):

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

As far as the central notion of barrier is concerned, I will assume that it is defined as follows (cf. Chomsky (1986), Baker (1988), Sportiche (1989), Cinque (1990), and Müller & Sternefeld (1990; 1993), among many others):[3]

(2) *Barrier*:
XP is a barrier for α iff there is an X^n ($0 \leq n \leq P$) such that (a), (b), and (c) hold:
a. X^n includes α.

---

[2] Minimize Chain Links (Relativized Minimality) is not to be confused with the Minimal Link Condition (MLC) of Chomsky (1995), which is basically a local version of the transderivational economy constraint Shortest Path Condition; see below. The MLC will become relevant for remnant movement in chapter 5, and I will address it then.

[3] This specific formulation is the one adopted in Müller & Sternefeld (1993), where further consequences of this notion of barrier are discussed. In order to serve as the basis of a comprehensive theory of locality, this definition might have to be slightly modified and extended. However, it may suffice for present purposes.

b. X^n is not a complement.
c. X^0 is distinct from Y^0, where XP is the complement of Y^0.

According to (2), any XP is a barrier if its head is distinct from the next higher head. Non-distinctness of heads resolves barrierhood; I will assume, essentially following Baker (1988), that non-distinctness can be achieved by overt (i.e., S-structure) or abstract (i.e., LF) incorporation; the latter is overtly signalled by co-indexing. However, given (2-b), such incorporation is confined to complements. Thus, XP is invariably a barrier (for something it includes) if it is a specifier or an adjunct, but not if it shows up in a complement position – in the latter case, barrierhood can be cancelled by non-distinctness of heads. Suppose now that V and the heads of its extended projection (AGR_O, T, AGR_S, etc., plus C) are always non-distinct, due to either overt incorporation or co-indexing at S-structure (i.e., abstract incorporation). Then, it follows that VP and the functional projections that are part of the extended projection of V (AGR_OP, TP, AGR_SP, etc.) are never barriers for overt movement (if they occur in situ). As for C, suppose that C is distinct from the embedding, governing head (normally V), at least when C is finite and in languages like German or English. Under these assumptions, a CP in complement position will be a barrier for anything included in C', but not for an item that appears in SpecC, given (2). If some category is included in C', CP is a barrier for it because C' is not a complement; however, if some category is included in CP, but not in C' (i.e., if it is in SpecC), then a complement CP is not a barrier for this category because in this case there is no projection of C that is a non-complement and that includes it (C' is a non-complement, but does not include it, and CP includes it, but is a complement, by assumption).

Along these lines, violations of the *Wh*-Island Condition are derived by the Barriers Condition, in more or less the same way as in Chomsky's (1986) theory of barriers; cf. (3):

(3) *How$_1$ did John tell you [$_{CP}$ when$_2$ to fix the car t$_2$ t$_1$] ?

Turning next to CPs in specifier positions, it follows from (2) that such items are always barriers and therefore block extraction, due to the Barriers Condition (1). This is shown for extraction from a finite subject CP in German in (4).[4,5]

---

[4] I assume here that the subject clause CP_2 occupies the SpecI position at the point where extraction takes place, despite the fact that it ends up in an extraposed position. I will address this issue in some detail in chapter 4.

[5] A well-known caveat is that subject infinitives in German do not block *wh*-extraction (or topicalization; they do block scrambling, though); cf., e.g., Haider (1983, sect. 2; 1993, 158-160), Sternefeld (1985a, sect. 3.3), and Grewendorf (1989, ch. 2.9) for discussion. An example that shows this is (i):

(i) Was$_1$ hat [$_{CP}$ t'_1 PRO t$_1$ zu beanstanden] sich nicht gehört ?
what$_{acc}$ has to object to not been proper

(4) *Wie$_1$ hat es ihn geärgert [$_{CP_2}$ t$'_1$ dass man nur t$_1$ nach Graz kommen
How has it him$_{acc}$ annoyed that one only to Graz get
kann] ?
is able to

Similarly, a CP in an adjoined position is a barrier, and *wh*-extraction from such a CP invariably violates the Barriers Condition. Thus, the illformedness of examples like (5-a) in English and (5-b) in German is accounted for.

(5) a. *How$_1$ did you leave [$_{CP}$ before fixing the car t$_1$] ?
b. *Wie$_1$ bist du eingeschlafen [$_{CP}$ nachdem du das Auto t$_1$ repariert hast] ?
how are you fallen asleep after you the car fixed have

Next, it follows that NP is a barrier in a specifier position (subject NPs), whereas it is not a barrier in a complement position (object NPs), assuming that non-distinctness of heads can be achieved via abstract incorporation. This is shown for German in (6).[6]

(6) a. *[$_{PP}$ Über wen]$_1$ hat [$_{NP}$ ein Buch t$_1$] den Karl beeindruckt ?
about whom has a book$_{nom}$ ART Karl$_{acc}$ impressed
b. [$_{PP}$ Über wen]$_1$ hat der Karl [$_{NP}$ ein Buch t$_1$] geschrieben ?
about whom has ART Karl$_{nom}$ a book$_{acc}$ written

Since there can be no movement to complement position (cf. Chomsky (1981; 1995)), it also follows that movement of an XP invariably turns XP into a barrier according to the notion in (2) (this will become important in the discussion of Freezing and Anti-Freezing effects in chapter 4).

Given the overall framework, the question that has to be asked is this: Is the Barriers Condition ‘minimalist’ in spirit? Arguably, it is not, and ideally, one would hope that its effects can be derived from some deeper assumptions in one way or another. In fact, a derivation of (1) from more minimalist principles does not seem to be impos-

According to the definition of barrier in (2), the embedded CP in (i) should block *wh*-extraction, an incorrect result. Indeed, the correct generalization about extraction from infinitives in German seems to be that *wh*-movement and topicalization from an argument infinitive are always possible, irrespective of the grammatical function of the clause and irrespective of where the clause appears in the middle field (i.e., whether it is scrambled or not). I think that these facts ultimately demand a radically different concept of barrier (in terms of minimality, not in terms of (non-) complementhood). However, for present purposes, I will abstract away from this complication and continue to assume the more common notion of barrier in (2); see Müller (1995, 54-56 & 75-76) for an approach to extraction from infinitives that relies on a different notion of barrier.

[6]Note that many further restrictions on extraction from NP are still unaccounted for under these assumptions. Thus, we have seen in chapter 1 that extraction from NP also depends on certain lexical properties of the governing verb, is subject to the Specified Subject Condition, etc. Ideally, all these further restrictions should fall out of the theory of barriers assumed here (and I think they do), but I will not go into that; cf. Müller (1995, ch. 2).

sible (cf. in particular Kitahara (1994a) and Takahashi (1994)), but this is beyond the scope of the present study. That said, let me now address economy constraints.

3. Economy Constraints

Formally, there are two types of economy constraints. On the one hand, there are *transderivational* (or *global*) economy constraints like Procrastinate, the Fewest Steps condition, and the Shortest Path Condition; on the other hand, there is the simple *derivational* (or *local*) economy constraint Last Resort. The basic difference between these two types of economy constraints is this: In order to determine whether or not a transderivational constraint is fulfilled by a derivation (or the resulting pair of LF and PF representations), one has to compare several derivations. In contrast, in order to determine whether or not a derivational constraint is fulfilled by a derivation (or the sentence it produces), one only has to look at single steps in this derivation. Let me start with the transderivational constraint Procrastinate.

3.1. Procrastinate

In addition to the classic notion of (un-)grammaticality, Chomsky (1993) introduces the notion of *convergence* (and the complementary notion of *crash*), which can be defined as in (7).

(7) *Convergence*:
A derivation *converges* at LF (PF) iff it contains only legitimate LF (PF) objects (otherwise it *crashes*).

I will assume that a derivation that does not converge (i.e., crashes) at, say, LF always creates an LF representation that is "ill formed" or "ungrammatical." Nevertheless, the notions of "non-convergence" and "ungrammaticality" are not equivalent. The reason is that ungrammaticality may arise despite the fact that the derivation converges. In other words, even if a derivation creates two representations (at PF and LF) that contain only *legitimate* (PF and LF) objects, this derivation may lead to illformedness (ungrammaticality) – either because a local derivational constraint is violated or because a representational constraint applying at PF or LF is violated.

Clearly, then, the question arises as to what legitimate LF and PF objects are. Chomsky (1993; 1995) proposes various requirements for LF objects. For example, an NP is a legitimate LF object only if its Case is checked. Another assumption that I will make here (deviating somewhat from Chomsky (1995)) is that a *wh*-phrase is a legitimate LF object only if it binds a variable, for reasons of semantic interpretation (see Karttunen (1977), among many others):

(8) *Wh-Legitimacy*:
A *wh*-phrase is a legitimate LF object only if it binds a variable.

This implies that all *wh*-phrases that are in situ at S-structure are raised at LF. As far as PF is concerned, Chomsky (1993) suggests that one of the conditions regulating the legitimacy of PF objects is the following:

(9) Strong features are not legitimate PF objects.

Therefore, all strong features that appear on lexical items must be removed before S-structure (Spell-Out).[7] Removal of a feature that shows up on a lexical item X is accomplished by movement – either of X (or, by percolation, XP) to a target head with matching features, or, if X itself is the target head, by movement of Y or YP with matching features. Thus, feature checking is a two-place relation.[8] Features that are checked disappear, by assumption.[9] Furthermore, suppose that it is often not an inherent property of features to be strong or weak. For the sake of concreteness, I will assume that the feature [+wh] of C is strong in languages that exhibit overt *wh*-movement, like English and German, and weak in *wh*-in situ languages like Korean and Japanese (but see Watanabe (1992) for a different view):

(10) *Strength of [+wh]*:
The [+wh] feature of C is strong in English and German, and weak in Korean and Japanese.

The assumptions in (9) and (10) then suffice to explain why overt *wh*-movement is obligatory in *wh*-clauses in English:

(11) a. *I wonder [$_{\text{CP}}$ – C$_{[}$+wh] you saw who$_1$]
b. I wonder [$_{\text{CP}}$ who$_1$ C you saw t$_1$]

In the derivation that generates the surface structure representation (11-a), the strong [+wh] feature on the embedded C node has not been removed by syntactic movement before Spell-Out. Hence, it remains in the derivation until PF and thereby causes the derivation to crash, yielding illformedness. In contrast, in the derivation generating

---

[7]Chomsky (1993; 1995) suggests that the level of S-structure should be dispensed with and replaced by the operation Spell-Out that creates the branching of the derivation to PF and LF in the traditional T-model. Henceforth, I will often use the notions of "S-structure" and "Spell-Out" interchangeably – i.e., the use of the notion "S-structure" should not be taken to imply that there is such a level of representation.

[8]The "matching" features that participate in checking must be similar, but strictly speaking they cannot be identical. For instance, *wh*-movement to SpecC involves checking of the feature [+wh] that must show up on both C and the *wh*-phrase. However, neither does a feature [+wh] on C turn CP into a *wh*-phrase that can undergo syntactic *wh*-movement, nor does a feature [+wh] on a *wh*-phrase turn this *wh*-phrase into a possible landing site for *wh*-movement. More generally, we have to distinguish between a "target" feature F_T and a "movement" feature F_M. In what follows, I will implicitly presuppose this distinction.

[9]In Chomsky (1995), this idea appears in a slightly modified form. There, it is assumed that certain kinds of features (those that are "[+Interpretable]") are not actually removed by checking, whereas other kinds of features (those that are "[–Interpretable]") are. I will come back to this issue.

(11-b), the strong [+wh] feature on C has been removed by overt *wh*-movement; consequently, the derivation converges at PF, and the sentence is well formed.

Furthermore, the assumptions in (9) and (10) explain why the *wh*-phrase *wae* ('why') may stay in situ in the Korean example (12-a). However, it is not yet clear why it *has* to stay in situ and cannot undergo *wh*-movement to the specifier of the [+wh] SpecC position before Spell-Out, as in (12-b).[10]

(12) a. [$_{CP}$ Nŏ-nŭn [$_{CP}$ Ch'ŏlsu-ka wae$_1$ o-ass-ta-ko] saengkakha-ni] ?
you-top Ch'ŏlsu-nom why come-IMP-DECL-C believe-Q
'Why$_1$ do you think Ch'ŏlsu came t$_1$?'

b. *[$_{CP}$ Wae$_1$ nŏ-nŭn [$_{CP}$ t$'_1$ Ch'ŏlsu-ka t$_1$ o-ass-ta-ko] saengkakha-ni] ?
why you-top Ch'ŏlsu-nom come-IMP-DECL-C believe-Q

What happened in (12-b) is that the weak [+wh] feature on the matrix C node has been checked and removed before Spell-Out by *wh*-movement of *wae*. This derivation converges, just like the one generating (12-a) does, so the illformedness of (12-b) cannot possibly follow from the presence of some illegitimate object. However, the illformedness follows from Procrastinate. Informally, Procrastinate can be stated as follows (cf. Chomsky (1993, 30)):

(13) *Procrastinate* (informal version):
LF movement is "cheaper" than overt movement.

A more precise statement of Procrastinate immediately reveals that it is a transderivational constraint (see, e.g., Marantz (1995, 357)):

(14) *Procrastinate*:
If two derivations D_1 and D_2 are in the same reference set, and D_1 differs from D_2 in that an item α is moved covertly (after Spell-Out) in D_1 and overtly (before Spell-Out) in D_2, then D_1 is to be preferred over D_2.

The reference set determines which derivations compete with respect to transderivational economy constraints. For present purposes, it can be defined as follows:

---

[10] I have chosen examples involving long-distance *wh*-dependencies in Korean to make sure that the evidence is not blurred by the phenomenon of *wh*-scrambling, which at first sight could pass as overt *wh*-movement. However, *wh*-scrambling of adjuncts, like all adjunct scrambling in Korean (and unlike *wh*-movement in general), is clause-bound. Hence, (12-b) cannot be generated by scrambling, and the illformedness of the example shows that it cannot be generated by overt *wh*-movement either, which is the point that this sentence is supposed to make.

(15) *Reference Set* (Chomsky (1995, 227)):
Two derivations D_1 and D_2 are in the same reference set iff they start with the same numeration and converge at LF and PF.

All the lexical items that are used in a derivation are drawn from the lexicon before computation starts; this array of lexical items (which includes n occurrences of an item X if X is to be used n times in the derivation) is called the *numeration*.

The two derivations D_1 and D_2, associated with the two Korean sentences in (12-a) and (12-b) respectively, are in the same reference set; hence, Procrastinate prefers D_1 to D_2. In contrast, the two derivations D_1 and D_2 associated with the two English sentences in (11-a) and (11-b) are *not* in the same reference set – D_1, with the S-structure representation in (11-a), does not converge (it contains a strong [+wh] feature as an illegitimate PF object), and therefore does not compete with D_2 (which generates the S-structure representation (11-b)). Hence, (11-b) in English is not blocked by (11-a), whereas (12-b) in Korean is blocked by (12-a).

Several other applications of the Procrastinate principle have been suggested in the literature, but for the time being, I will leave it at that and turn to the Fewest Steps condition.

3.2. Fewest Steps

The Fewest Steps condition was first proposed in Chomsky (1991) and can, in a way, be viewed as the most basic economy constraint. Like Procrastinate, it is transderivational. It can be defined as in (16) (cf. Chomsky (1991; 1993; 1995)):

(16) *Fewest Steps*:
If two derivations D_1 and D_2 are in the same reference set and D_1 involves fewer operations than D_2, then D_1 is to be preferred over D_2.

Thus, the Fewest Steps condition requires that of two competing derivations, the one that involves fewer syntactic operations (like movement) must be chosen. Since the formulations of Procrastinate in (14) and of the Fewest Steps condition in (16) are similar, one might hope that the former constraint can in fact be derived from the latter (apparently more general) one. However, as it stands, Procrastinate cannot be derived from Fewest Steps. To see this, let us consider again the ill-formed Korean example in (12-b) and try to figure out whether it (more precisely, the derivation that generates it) is also excluded by the Fewest Steps condition. The answer is no: Whether the *wh*-adjunct *wae* moves before or after Spell-Out does not appear to have any consequences with regard to the number of movement operations involved. Thus, from the point of view of Fewest Steps, the derivation D_1 generating (12-b) is just as costly as the derivation D_2 generating (12-a). Therefore, Fewest Steps does not discriminate between the two derivations – in order to block D_1 by D_2, it seems that an appeal to (something like) Procrastinate is necessary. More generally, then, unless one can find an independent reason why one and the same movement operation involves more steps before Spell-Out than after Spell-Out, we have to conclude that the effects of Procrastinate cannot be derived from the Fewest Steps condition.[11]

That said, let me now illustrate the Fewest Steps condition by discussing five of its consequences.

3.2.1. Illicit Wh-Topicalization in English

Languages like English exhibit a strict prohibition against topicalization of *wh*-phrases (cf. Lasnik & Uriagereka (1988), Lasnik & Saito (1992), and Rizzi (1996), among others, and subsection 5.2.1 below for evidence from German). Consider the following examples:

(17) a. Who$_1$ t$_1$ said [$_{CP}$ that [$_{IP}$ John likes who$_2$]] ?
b. Who$_1$ t$_1$ said [$_{CP}$ that Mary$_2$ [$_{IP}$ John likes t$_2$]] ?
c. *Who$_1$ t$_1$ said [$_{CP}$ that who$_2$ [$_{IP}$ John likes t$_2$]] ?

First, note that the embedded CP in (17-a) is a [–wh] clause. Therefore, the sentence must be interpreted as a multiple question, with *who*$_2$ in situ at S-structure but (by assumption) in the matrix SpecC position at LF; the example is well-formed. Turning next to (17-b), it is evident that there is no general prohibition against topicalization in an embedded CP with a matrix verb like *say* and in the presence of overt *wh*-movement in the matrix clause; the topicalized item here is the [–wh] phrase *Mary*$_2$. The interesting thing now is that such a topicalization operation apparently is not permitted if the moved item is a *wh*-phrase, as in (17-c). At first sight, this ban on *wh*-topicalization is surprising. However, as noted by Epstein (1992), it is explained directly by the Fewest Steps condition in (16), given that *wh*-in situ phrases undergo LF raising. The basic idea is that (17-c) is blocked by (17-a) because the derivations generating the two S-structure representations compete and the derivation that yields (17-a) is more economical (i.e., it involves fewer operations). More precisely, the argument goes as follows.

Given that all *wh*-phrases must occupy a SpecC (substitution or adjoined) position at LF, (17-a) and (17-c) both end up with the following LF representation:[12]

(18) *LF representation of (17-a) and (17-c)*:
[$_{SpecC}$ [$_{SpecC}$ who$_1$] who$_2$] t$_1$ said [$_{CP}$ that [$_{IP}$ John likes t$_2$]] ?

---

[11] Obviously, if such an independent reason is contrived, things might look different. Thus, Kitahara (1995, 67-74) suggests that LF movement of some category α to some position β may involve one operation fewer than syntactic movement of α to β does. Along these lines, the effects of Procrastinate can indeed be derived from the Fewest Steps condition, and the Procrastinate principle can be dispensed with.

[12] For the time being, I assume, following Rudin (1988), that adjunction to SpecC must be to the right, but nothings depends on this assumption in the present context. See, however, chapter 4 for a modification of this view.

Moreover, suppose that both derivations converge (this implies, in particular, that neither derivation results in a PF representation that contains strong features). Being equipped with the same numeration, the derivations D_1 (with the S-structure representation in (17-a) and the LF representation in (18)) and D_2 (with the S-structure representation in (17-c) and the LF representation in (18)) are then in the same reference set and compete with respect to transderivational economy constraints. However, D_1 involves one movement operation applying to *who*$_2$ (only *wh*-movement at LF), whereas D_2 involves two movement operations applying to *who*$_2$ (topicalization at S-structure followed by *wh*-movement at LF). Hence, D_1 is preferred over D_2, and (17-c) is ruled out by Fewest Steps.

3.2.2. Illicit Wh-Scrambling in German

The well-known prohibition against *wh*-scrambling in German (see Fanselow (1990), Stechow & Sternefeld (1988), Rizzi (1996), and Müller & Sternefeld (1993; 1996)) can be derived along the same lines. The sentence in (19-a) is a well-formed multiple question with a direct *wh*-object *was* ('what') in situ. (19-b) shows that direct objects can be scrambled in front of the subject in *wh*-clauses without inducing ungrammaticality. However, such scrambling is not possible if the moved item is not a [–wh] phrase (as in (19-b)), but a [+wh] phrase (as in (19-c)).

(19) a. Wie$_1$ hat [$_{IP}$ der Fritz t$_1$ was$_2$ repariert] ?
how has ART Fritz what$_{acc}$ fixed

b. Wie$_1$ hat [$_{IP}$ [das Auto]$_2$ [$_{IP}$ der Fritz t$_1$ t$_2$ repariert]] ?
how has the car$_{acc}$ ART Fritz fixed

c. *Wie$_1$ hat [$_{IP}$ was$_2$ [$_{IP}$ der Fritz t$_1$ t$_2$ repariert]] ?
how has what$_{acc}$ ART Fritz fixed

Again, the illformedness of (19-c) follows from the Fewest Steps condition. The reason is that both (19-a) and (19-c) have an LF representation like (20).

(20) [$_{SpecC}$ [$_{SpecC}$ wie$_1$] was$_2$] [$_{IP}$ der Fritz t$_1$ t$_2$ repariert hat]

Given that the derivations D_1 (generating (19-a)) and D_2 (generating (19-c)) have the same numeration and converge, it follows that D_1 and D_2 are in the same reference set. This, in turn, means that D_1 blocks D_2: D_1 involves one application of movement to the *wh*-phrase *was* ('what') (viz., only LF raising to SpecC), whereas D_2 reaches the LF representation in (20) by two applications of movement to *was* (viz., scrambling before Spell-Out followed by LF movement to SpecC).[13]

[13]It appears that *wh*-scrambling in German is not completely impossible under certain circumstances. For one thing, there is a clear connectedness effect (in the sense of Kayne (1984)), i.e., by adding more *wh*-phrases in a clause, scrambling of a *wh*-phrase tends to become more tolerable. In addition, *wh*-scrambling leads to better results if the *wh*-phrase is D-linked (cf. Pesetsky (1987)), as has been noted by Fanselow

Intuitively, the Fewest Steps condition rules out *wh*-topicalization and *wh*-scrambling because these movements are superfluous – the resulting LF representations can be achieved in a simpler way, without optional topicalization and optional scrambling. Although this result is clearly desirable in cases of ill-formed examples like (17-c) and (19-c), it raises an important question: Why is optional movement ever permitted under Fewest Steps? I address this issue in the following subsection.

3.2.3. Optional Movement

The following examples show that (embedded) topicalization is optional in English:

(21) a. I think [$_{\text{CP}}$ that John likes Mary$_1$]
b. I think [$_{\text{CP}}$ that Mary$_1$ John likes t$_1$]

At first sight, this might come as a surprise, given the Fewest Steps condition, because a derivation D$_1$ via overt topicalization (as in (21-b)) involves one more movement operation than a derivation D$_2$ without topicalization (as in (21-a)).

The same problem is posed by scrambling. (22) exhibits the optionality of scrambling in German, and (23) shows the same for Korean.

(22) a. dass [$_{\text{IP}}$ keiner den Fritz$_1$ gesehen hat]
that no-one$_{nom}$ ART Fritz$_{acc}$ seen has
b. dass [$_{\text{IP}}$ den Fritz$_1$ [$_{\text{IP}}$ keiner t$_1$ gesehen hat]]
that ART Fritz$_{acc}$ no-one$_{nom}$ seen has

(23) a. Ch'ŏlsu-ka [$_{\text{VP}}$ Sunhi-eke i ch'aek-ŭl chu-ŏss-ta]
Ch'ŏlsu$_{nom}$ Sunhi$_{dat}$ this book$_{acc}$ gave
b. Ch'ŏlsu-ka [$_{\text{VP}}$ i ch'aek-ŭl$_1$ [$_{V\text{P}}$ Sunhi-eke t$_1$ chu-ŏss-ta]]
Ch'ŏlsu$_{nom}$ this book$_{acc}$ Sunhi$_{dat}$ gave

The potential problem for the Fewest Steps condition which is posed by apparently optional types of movement has often been noted in the literature. One possible solution that Fukui (1993) has proposed for scrambling (also see Epstein (1992) for relevant remarks) is to assume that movement operations which are sufficiently "local" simply do not count as an operation for the purposes of Fewest Steps. Fewest Steps, according to this view, "overlooks" small movements like scrambling and pays attention only to "more visible" movements like *wh*-movement. But this view is problematic – we have just seen that in order to rule out *wh*-topicalization as in (17-c) and *wh*-scrambling as

---

(1992). Whereas the nature of the connectedness effect must be left open here, it can be noted that the D-linking effect would follow under Pesetsky's assumption that D-linked *wh*-phrases are not operators (i.e., lack the relevant feature), and hence, do not have to undergo LF raising to SpecC: Given that subsequent LF raising does not have to take place, *wh*-phrases are predicted to behave like [–wh] phrases with respect to the issue of scrambling – an issue to which I will turn in the next subsection.

in (19-c) by Fewest Steps, even “small” movements like embedded topicalization in English and scrambling in German must count as syntactic operations for the purposes of Fewest Steps.

For that reason, I would like to pursue another approach to the optionality of certain movement operations and contend that this optionality is indeed only apparent. Thus, suppose that topicalization in English and scrambling in German and Korean are in fact not optional movement types. Rather, in these cases, a strong feature is present that must be checked and thereby eliminated by movement before Spell-Out. On this view, what is optional about these movement types is not the movement operation per se, but rather the presence of the triggering feature.

For the sake of concreteness, I will make the following assumptions about the movement types *wh*-movement, topicalization, and scrambling. First, *wh*-movement is movement to SpecC – either substitution or, if SpecC is already filled (as in many cases of LF *wh*-movement), adjunction; overt *wh*-movement is triggered by a strong feature [+wh] on C. Second, topicalization is substitution in SpecTop (Sternefeld (1991), Müller & Sternefeld (1993; 1996), Müller (1995), Zwart (1993)), where TopP intervenes between CP and IP, and Top is the landing site of V/2 movement in the Germanic languages; topicalization is triggered by a strong [top] feature on Top.

Under this view, typical instances of (embedded) topicalization in the Germanic languages look as (24-a) in English, (24-b) in Danish (cf. Vikner (1995)), and (24-c) in German:[14]

(24) a. I think $[_{\text{CP}}$ – $[_{\text{C}}$ that] $[_{\text{TopP}}$ $[_{\text{NP}}$ Mary $]_1$ $[_{\text{Top}}$ –] $[_{\text{IP}}$ John likes t_1]]]

b. Hun sagde $[_{\text{CP}}$ – $[_{\text{C}}$ at] $[_{\text{TopP}}$ $[_{\text{NP}}$ denne bog $]_1$ $[_{\text{Top}}$ har] $[_{\text{IP}}$ Bo ikke
she said that this book has Bo not
læst t_1]]]
read

c. Sie sagte $[_{\text{CP}}$ – [C –] [TopP [NP dieses Buch $]_1$ $[_{\text{Top}}$ habe] $[_{\text{IP}}$ sie
she said this book_{acc} has_{subj} she_{nom}
nie t_1 gelesen]]]
never read

Finally, I assume that scrambling in German and Korean is left-adjunction to some XP in the middle field (Fanselow (1990), Webelhuth (1992), Saito (1985)), where X

---

[14] Given that topicalization in English, Danish, and German is viewed as the same kind of phenomenon and analyzed in terms of substitution in SpecTop, something has to be said about the obvious differences between the topicalization operations in these languages. Thus, embedded topicalization in English regularly occurs with a complementizer *that* present (cf. Grimshaw (1994), but also Rochemont (1989) for a qualification), but (usually) without V/2 movement (Subject-Aux Inversion) (see Rizzi (1996)); embedded topicalization in Danish systematically exhibits a co-occurrence of complementizer and V/2; embedded topicalization in German, in contrast, requires V/2 movement, but a complementizer is not permitted. See Müller & Sternefeld (1993) for an approach to these differences within the unified analysis of topicalization as substitution in SpecTop.

$=$ V or I.[15] This movement is triggered by a strong [scr] feature on X, as is shown schematically in (25) (note that this implies that lexical categories like V can be the target of movement and check off features of the moved item).

(25) ... $[_{\mathrm{XP}}\ \alpha_1\ [_{\mathrm{XP}}$ ... t_1 ...]] ...

Given the well-known semantic and functional heterogeneity of topicalization in German (cf., e.g., Grewendorf (1989)) and in English (cf., e.g., Culicover (1991)), and given that there is no unique semantic operation that could be attributed to scrambling in German, [scr] and [top] are, to my mind, best viewed as purely formal features without any inherent semantic or functional content. Accordingly, concepts like “LF scrambling” and “LF topicalization” appear poorly motivated. Hence, I will assume that [scr] and [top], being semantically empty, are always strong, in the sense that they induce *overt* movement; cf. Chomsky (1995, 349-355). Notice that this view does not rule out the possibility that, e.g., scrambling may have semantic effects, as has often been argued (cf. Diesing (1992), de Hoop (1992), Büring (1993), and Meinunger (1995), among others). Rather, what is presupposed here is that in cases where scrambled and non-scrambled orders differ in truth conditions, this is not due to a specific semantic contribution of some feature triggering scrambling, but rather to the relative surface order of quantified items – a factor that is known to be relevant for semantic interpretation independently of movement. To sum up: My view is that [scr] and [top] are semantically empty (or [–Interpretable], see subsection 3.4.2 below) features, but the topicalization and scrambling operations triggered by these features may or may not have semantic effects, depending on the nature of the moved item (e.g., quantificational or not) and the nature of the material that is crossed by the movement operation.

Do these assumptions suffice to permit topicalization in English and scrambling in German and Korean, in accordance with the transderivational Fewest Steps condition in (16)? The answer depends on specific assumptions about numerations. Recall first the notion of reference set adopted so far (cf. (15)).

(26) *Reference Set*:
Two derivations D_1 and D_2 are in the same reference set iff they start with the same numeration and converge at LF and PF.

Now, if we assume that two numerations are not identical if they differ in the feature specification on some head (like Top), the Fewest Steps condition predicts that derivations involving topicalization or scrambling are not blocked by (otherwise identical)

[15] Here and henceforth, I will assume a phrase structure for German sentences that does without the articulated structure that is postulated under the split INFL hypothesis (see Pollock (1989), Chomsky (1991; 1993; 1995)); i.e., I postulate that AGR_O, T, and AGR_S are merged into a single I node. At present, this is mainly to simplify exposition; however, I will present independent (though theory-internal) evidence for this move in chapters 4 and 5. Also cf. the pertinent remarks in Chomsky (1995, sect. 10).

derivations where these movements have not occurred, as intended. However, if we were to assume that two numerations are identical if they differ only in the feature specification on some head (like Top), the Fewest Steps condition still rules out "optional" movement like topicalization and scrambling. Thus, initially there appears to be reason to adopt the first hypothesis over the second – only under the first hypothesis does it follow that a derivation D_1 with embedded topicalization in English, as in (21-b) (= (24-a)), is not in the same reference set as the derivation D_2 generating the in situ counterpart (21-a), because a strong [top] feature is present on Top in D_1's numeration, but not in D_2's numeration. (A similar reasoning applies in the case of scrambling.)

But if this is so, Fewest Steps does not rule out (17-c) anymore, because the two derivations generating (27-a) (= (17-a)) and (27-b) (= (17-c)) then also have different numerations.

(27) a. Who$_1$ t$_1$ said [$_{CP}$ that [$_{TopP}$ [$_{IP}$ John likes who$_2$]]] ?
b. *Who$_1$ t$_1$ said [$_{CP}$ that [$_{TopP}$ who$_2$ [$_{IP}$ John likes t$_2$]]] ?

For this reason, I will adopt the second hypothesis concerning the identity of numerations for the time being (but see section 5 below). This hypothesis can be formulated as in (28):

(28) *Identity of Numerations*:
Two numerations are identical if they differ only in the feature specification on some head.

This implies that illicit *wh*-topicalization as in (17-c) and illicit *wh*-scrambling as in (19-c) can still be ruled out by Fewest Steps; but it also implies that the notion of reference set must be modified so as to permit topicalization and scrambling at all. Following ideas of Kitahara (1993, 113), Fox (1994, 32), Reinhart (1994, 17), and Grimshaw (1994, 3), I would like to suggest the notion of reference set in (29); here, the requirement is added that competing derivations must yield identical LF representations.

(29) *Reference Set* (revised):
Two derivations D_1 and D_2 are in the same reference set iff they start with the same numeration, converge at LF and PF, and yield the same LF representation.

This revised notion of reference set still suffices to rule out illicit *wh*-topicalization and *wh*-scrambling, as discussed in the preceding subsection; for instance, the derivations D_1 (generating (27-a)) and D_2 (generating (27-b)) yield identical LF representations and are therefore in the same reference set, so that Fewest Steps chooses D_1 over D_2, as desired. However, assuming that movement operations like scrambling and topicalization are not undone at LF, it follows that a derivation generating a scrambling structure as in the German example (22-b) is not in the same reference set as a derivation in which scrambling has not occurred, as in the German example (22-a); hence, (22-b) is not blocked by the 'more economical' (22-a). The same goes for topicalization, as in the English examples in (21-a) and (21-b): (21-b) is not blocked by (21-a) because the derivations generating these S-structure representations do not yield identical LF representations and are therefore not in the same reference set.[16]

In conclusion, Fewest Steps now rules out "optional" overt movement (conceived of as obligatory movement triggered by optional features) only if this movement is undone by further LF movement, but not if the effects of this movement persist at LF.

3.2.4. Form Chain vs. Move α

A fourth consequence of the Fewest Steps condition relates to the issue of successive cyclicity. The Barriers Condition forces successive-cyclic movement, as in (30):

(30) Why$_1$ do you think [$_{\text{CP}}$ t'_1 that Mary came t_1] ?

If the embedded SpecC position were skipped over in the course of long-distance movement, a locality violation should be expected, on a par with typical *wh*-island violations. Therefore, it was commonly postulated that an S-structure representation like (30) is derived by two applications of Move α – first to the embedded SpecC position, and then to the matrix SpecC position.

In contrast, it seems that Fewest Steps requires long-distance movement to apply in one fell swoop, in violation of the Barriers Condition, as in (31):

---

[16]Three remarks are in order here concerning the notion of "identical LF representation." First, I assume that intermediate traces do not induce LF-distinctness. That is, even though the LF representation of a sentence like (27-b) might contain an intermediate trace t'_2 in the embedded SpecTop position (but see subsection 3.2.5 below) while the LF representation of a sentence like (27-a) does not contain such an intermediate trace, the two LF representations still count as identical, in the sense of (29). Second, if string-vacuous scrambling of, e.g., an object NP to VP does not generate a different LF representation in this technical sense, this operation will be blocked by Fewest Steps, which is arguably a welcome result (cf. Ross (1967)). (Note that this would not rule out scrambling that is locally string-vacuous but does eventually result in a different LF representation, as in (3) from chapter 1.) Finally, it is worth keeping in mind that the notion of "identical LF representation" is not the same as the notion of "identical interpretation" (here I differ from, e.g., Fox (1994) and Reinhart (1994), who postulate that "identical interpretation" is the relevant notion for defining reference sets). Thus, although I presuppose here that scrambling as in (22-b) in German and (23-b) in Korean is not undone at LF and is therefore not blocked by the in situ variants in (22-a) and (23-a) (because they have a different LF representation), this does not necessarily imply that sentences like (22-a) and (22-b) actually differ in interpretation. More generally, whereas cases that exhibit an interpretational difference between a scrambled and a non-scrambled order are compatible with both the notions "identical interpretation" and "identical LF representation" in the present system, other cases that show a lack of interpretational difference between a scrambled and a non-scrambled order (as is arguably the case in, e.g., (22-a)/(22-b)) are compatible only with the latter notion. Similar considerations apply with respect to topicalization in languages like English, Danish, and German (cf. (24)). Assuming the concept of λ-conversion as a means of reconstruction (cf., for instance, Heycock & Kroch (1997) and Sternefeld (1997a) for articulated proposals), or assuming the copy theory (see chapter 3), sentences involving scrambling and topicalization structures at LF can still receive the same interpretation as their in situ versions.

(31) *Why$_1$ do you think [$_{\mathrm{CP}}$ – that Mary came t$_1$] ?

The reason is that a derivation D_1 that generates the S-structure representation in (30) involves two applications of Move α to the *wh*-phrase, as opposed to a derivation D_2 generating (31) that involves only one application of Move α to the *wh*-phrase. Since the fact that (31) does, and (30) does not, violate the Barriers Condition does not play a role for the question of whether D_1 and D_2 are in the same reference set and since, furthermore, both D_1 and D_2 converge (at LF and PF) and yield identical LF representations (in the sense explained in the last footnote), it seems that the Fewest Steps condition forces the choice of derivation D_2 generating (31), and the string should be ungrammatical, contrary to fact.

In order to solve this problem induced by the Fewest Steps condition, Chomsky (1991; 1993) proposes that successive-cyclic applications of Move α are to be viewed as parts of a single syntactic operation Form Chain. Form Chain can be defined as in (32):

(32) *Form Chain*:
Move α to its target position and insert intermediate traces in positions required by locality conditions.

Thus, (30) is not ruled out by Fewest Steps in favour of (31) anymore. Both derivations involve only one application of Form Chain to the *wh*-phrase *why* – the insertion of intermediate traces is costless.

3.2.5. Argument/Adjunct Asymmetries

Finally, a fifth interesting consequence of the Fewest Steps condition is that it contributes significantly to the derivation of argument/adjunct asymmetries with long-distance movement (cf. Chomsky & Lasnik (1993), on which the following subsection is based).

Locality violations are typically strong with adjuncts and more variable with arguments (see Lasnik & Saito (1984; 1992), Chomsky (1986), Rizzi (1990), Cinque (1990), and Manzini (1992), among others).[17] For example, if the Barriers Condition is violated in the course of adjunct movement, strong ungrammaticality arises (cf. (33-a)), but a locality violation with arguments is typically much weaker (cf. (33-b)).

(33) a. *How$_3$ do you wonder [$_{\mathrm{CP}_1}$ whether John said [$_{\mathrm{CP}_2}$ t'_3 C Mary solved the problem t$_3$]] ?

---

[17] In what one might call the "standard approach," Lasnik & Saito (1984) account for this asymmetry by assuming that adjunct traces cannot be deleted on the way to LF, which eventually follows from the stipulation that proper government of these traces (encoded in the assignment of a feature γ) is checked only at LF.

b. ??Who$_3$ do you wonder [$_{CP_1}$ whether John said [$_{CP_2}$ t'_3 C t_3 solved the problem]] ?

In both (33-a) and (33-b), t'_3 is marked as ungrammatical, due to a *wh*-island effect which is reducible to the Barriers Condition (*whether* making SpecC inaccessible as an escape hatch, cf. Chomsky (1986)). Chomsky & Lasnik (1993) suggest that the "offending" trace, i.e., the trace that is responsible for the locality violation, is marked by a star: *t. Next, it is assumed that a starred trace that appears in the derivation induces weak (Subjacency-like) ungrammaticality of a sentence and that a starred trace that is present at LF induces strong (ECP-like) ungrammaticality. The idea is then that the offending intermediate trace t'_3 is deleted on the way to LF in (33-b) but not in (33-a), so that (33-a) turns out to be strongly (ECP-like) ungrammatical and (33-b) to be weakly (Subjacency-like) ungrammatical. To accomplish this result, something has to be said about the conditions under which traces that are generated during the derivation can or must be deleted. To this end, Chomsky & Lasnik (1993, 546) postulate that 'legitimate LF objects' must be of the type (34-a) or (34-b).[18]

(34) *Legitimate LF objects*:

a. Uniform chains of the type C = $<\alpha_1,...,\alpha_n>$, where all members are either in an A-bar position or in an A-position.

b. Operator-variable constructions of the type $<\alpha,\beta>$, where α is in an A-bar position and β heads a uniform chain.

The argument chain in (33-b) C_1 = $<who_3,t'_3,t_3>$ is not a uniform chain (the initial trace t_3 is in an A-position, and the other members of C_1 are in A-bar positions), nor is it an operator-variable construction (the operator *who* is in an A-bar position, but t'_3 does not head a uniform chain). Hence, assuming that every legitimate LF object must meet at least one of the two conditions of (34), syntactic operations must apply on the way to LF in order to turn C_1 into a legitimate LF object. Deletion of t'_3 produces a legitimate operator-variable construction, with *who* in an A-bar operator position and t_3 heading a uniform (trivial) A-chain. Is this additional deletion operation permitted by Fewest Steps? It is: If deletion of the intermediate trace t'_3 does not apply on the way to LF in (33-b), the derivation does not converge at LF (yielding an illegitimate LF object). Given the notion of reference set in (29), a derivation D_1 in which t'_3 is deleted after Spell-Out in (33-b) is not blocked by the Fewest Steps condition in favour of a more economical derivation D_2 in which t'_3 remains undeleted (leading to non-convergence at LF), because only converging derivations can compete. Along these lines, the "standard" stipulation put forward in Lasnik & Saito (1984; 1992) and Chomsky (1986) concerning the deletability of intermediate argument traces can be derived.

[18] As noted above, there are additional conditions for LF legitimacy (NPs must have their Case features checked, *wh*-phrases must bind variables, etc.), but these are not relevant in the present context.

Turning now to adjunct movement in (33-a), it emerges that the adjunct chain C_2 = $<how_3$,t'_3,$t_3>$ already is a uniform chain (all chain members occupy A-bar positions) and hence a legitimate LF object. Therefore, deletion of t'_3 does not have to apply in order for convergence at LF to result. What is more, given Fewest Steps, deletion of the intermediate trace t'_3 *must* not apply: A derivation D_1 without deletion of t'_3 in (33-a) is in the same reference set as a derivation D_2 that differs from D_1 only in that t'_3 is LF-deleted; hence, Fewest Steps forces the choice of D_1 over D_2, even though D_1 contains an "offending" trace creating ungrammaticality while D_2 does not.[19]

Summarizing the last five subsections, the transderivational constraint Fewest Steps produces interesting and welcome consequences in some domains (concerning illicit *wh*-topicalization and *wh*-scrambling, and the LF-deletability of intermediate traces), and consequences that at first sight look somewhat less appealing in other domains (concerning the very existence of optional movement and the concept of Form Chain).

3.3. The Shortest Path Condition

The third (and final) transderivational economy constraint to be discussed here is the Shortest Path Condition (see Chomsky (1993; 1995), Kitahara (1993), Collins (1994), and Ackema & Neeleman (1995)). It can be formulated as follows (cf. Collins (1994, 56)):

(35) *Shortest Path Condition*:
If two derivations D_1 and D_2 are in the same reference set and the movement paths of D_1 are shorter than the movement paths of D_2, then D_1 is to be preferred over D_2.

Stated informally, the Shortest Path Condition says that if two derivations are in competition, then the one whose movement paths are shorter must be chosen. Length of movement paths can readily be defined in terms of dominating (Collins (1994, 56)) or c-commanding (Sternefeld (1997, 84)) intervening nodes crossed. Here, I will confine

---

[19]Note that under this assumption, there must be an intermediate object trace t'_1 in an A-bar position between *which car*$_1$ and t_1 in (i-a) – otherwise, there would be an initial trace *t_1 at LF ($<$*which car*,$t_1>$ would be a legitimate LF object, according to clause (b) of (34)), and (i-a) should be as ill formed as (i-b):

(i) a. ??Which car$_1$ did John tell you [$_{CP}$ how$_2$ (t'_1) to fix t_1 t_2] ?
b. *How$_1$ did John tell you [$_{CP}$ when$_2$ to fix the car t_2 t_1] ?

It is not quite clear where t'_1 can be located in (i-a), especially in the light of what will be said about improper movement in section 4 below. For the time being, I will assume, essentially following Müller & Sternefeld (1990; 1993), that t'_1 in (i-a) is in the SpecTop position of the embedded clause, an option that is confined to infinitival clauses (where SpecTop is underspecified and Top does not project inherent features). Alternatively, one might postulate that the projection of T(ense) provides another specifier as an escape hatch for A-bar movement; cf. Sabel (1995, ch. 2). See below.

myself to an intuitive understanding of the length of movement paths (see section 7 of chapter 3 for a definition) and will highlight two applications of the Shortest Path Condition.

3.3.1. The Shortest Path Condition and Yo-Yo Movement

A first consequence of the Shortest Path Condition, pointed out by Collins (1994, 53ff), is that it explains why lowering followed by raising before Spell-Out (so-called Yo-Yo movement; cf. Lasnik & Saito (1992)) is not possible. Several languages exhibit a diagnostic for the successive-cyclic nature of A-bar movement; i.e., they provide clear overt evidence for intermediate traces in SpecC. Among these languages are Irish (see McCloskey (1979)), Spanish (see Torrego (1984)), and Ewe, as discussed in Collins (1993; 1994). Relevant data from Ewe are given in (36).

(36) Kofi$_1$ e me gble [$_{CP}$ t$'_1$ be é/wo fo t$_1$]
Kofi FOC I said that he hit
'It was Kofi that I said that he hit.'

Here the NP *Kofi* has undergone long-distance focus movement; optionally, the subject pronoun 'he' may change its form from the regular *é* to *wo*. This optional morphological alternation is triggered by successive-cyclic movement through SpecC (i.e., by an intermediate trace in that position). Why exactly the complementizer *be* ('that') tolerates the form *wo* of the subject pronoun in SpecI only if its specifier position SpecC is filled, and why this change is optional and not obligatory, will not concern me here (see Collins (1993, ch. 4) for some discussion). What is relevant in the present context is simply the fact as such: Only if the specifier of the complementizer *be* is filled can *wo* show up in SpecI.

Not surprisingly, from a pre-theoretic point of view at least, focus movement that originates in the matrix clause and targets a higher position does not trigger this morphological change in the subject position of the embedded clause. This is shown by the contrast between (37-a) (with the regular subject pronoun *é*) and (37-b) (with the form *wo*).

(37) a. Kofi$_1$ e me gble na t$_1$ [$_{CP}$ be é fo Kosi]
Kofi FOC I said to that he hit Kosi
'It was Kofi that I told that he hit Kosi.'
b. *Kofi$_1$ e me gble na t$_1$ [$_{CP}$ be wo fo Kosi]
Kofi FOC I said to that he hit Kosi
'It was Kofi that I told that he hit Kosi.'

However, upon closer inspection, things are not that simple. A derivation of (37-b) must be ruled out in which *Kofi* moves to its ultimate position in the matrix clause via an intermediate step of lowering into the embedded SpecC position, where it triggers the morphological change. Various conditions come to mind that might rule out such a derivation,[20] but it is at least worth noting that a derivation in which *Kofi* undergoes Yo-Yo movement, as in (38), is excluded by the Shortest Path Condition.

(38) *Kofi$_1$ e me gble na t$_1$ [$_{CP}$ t$'_1$ be wo fo Kosi]

The movement path of *Kofi*$_1$ is longer in derivation D$_2$ via (38) than in derivation D$_1$ via (37-b). Hence, D$_1$ is to be preferred over D$_2$, according to the Shortest Path Condition (35); but in this derivation, the requirement that *wo* appears only if the local SpecC position is filled is violated.[21] More generally, the Shortest Path Condition accounts for the fact that an XP generated in a higher clause can never induce a reflex of successive-cyclic movement in a lower clause.

3.3.2. The Shortest Path Condition and Superiority Effects

A second consequence of the Shortest Path Condition concerns Superiority effects (cf. Chomsky (1973)). A typical pair of examples exemplifying the Superiority effect in English is given in (39).

(39) a. I wonder [$_{CP}$ who$_1$ C [$_{IP}$ t$_1$ bought what$_2$]]
b. *I wonder [$_{CP}$ what$_2$ C [$_{IP}$ who$_1$ bought t$_2$]]

According to the classic Superiority condition, (39-b) is ungrammatical because the highest *wh*-phrase in the clause (the subject *who*) has failed to undergo overt movement to SpecC; rather, the lower *wh*-object *what* is moved to SpecC. Basically the same holds for the examples in (40), with the difference that in this case, two *wh*-objects are involved (rather than a subject and an object).

(40) a. Whom$_1$ did John persuade t$_1$ [$_{CP}$ to visit whom$_2$] ?
b. *Whom$_2$ did John persuade whom$_1$ [$_{CP}$ t$'_2$ to visit t$_2$] ?

Again, the Superiority condition blocks overt movement of the lower *wh*-phrase *whom*$_2$, as in (40-b), in favour of overt movement of the higher *wh*-phrase *whom*$_1$, as in (40-a).

The basic idea behind the Superiority condition is maintained under an approach in terms of the Shortest Path Condition, as suggested in Chomsky (1993, 14) and developed in full in Kitahara (1993, 113ff). According to the Shortest Path Condition

---

[20]See chapter 3 below for discussion.

[21]Two remarks concerning the presentation of this argument are in order. First, it must be ensured that D$_1$ and D$_2$ are in the same reference set even though the embedded subject pronoun takes a different form; cf. Collins (1994) for relevant remarks. And second, Collins (1994, 54) actually assumes that lowering in (38) is preceded by adjunction of *Kofi*$_1$ to the matrix VP, for reasons of locality. Assuming that locality constraints are not violated by lowering from t$_1$ to t$'_1$ in (38), this additional step does not seem to be necessary. I will briefly return to this issue in subsection 4.2.1 of chapter 3.

in (35), the (a)-examples are preferred over the (b)-examples because the movement path with index 1 is shorter than the movement path with index 2.

So far, so good. However, it is clear that this solution rests on the assumption that *wh*-in situ phrases do not undergo *wh*-movement at LF. If they undergo LF movement, the two *wh*-movement paths of, e.g., (40-b) and (40-a) will have the same length. But the assumption that *wh*-movement at LF does not exist is incompatible with the Fewest Steps account of illicit *wh*-topicalization and illicit *wh*-scrambling discussed above (see subsection 3.2). Moreover, given the notion of reference set in (29), repeated here as (41), one would have to make sure that the LF representations of (40-b) and (40-a) are sufficiently "similar," so that the two derivations compete with respect to the Shortest Path Condition – under the assumption that *wh*-in situ phrases do not undergo LF movement, this result cannot be achieved straightforwardly (see Kitahara (1993), Grimshaw (1994), and Sternefeld (1997) for discussion of this issue).

(41) *Reference Set* (= (29)):
Two derivations D_1 and D_2 are in the same reference set iff they start with the same numeration, converge at LF and PF, and yield the same LF representation.

For these reasons, I conclude that accounting for Superiority effects in terms of the Shortest Path Condition is problematic. I will return to this issue (and will eventually adopt a local version of this constraint, the Minimial Link Condition (MLC), in chapter 5); for now, though, I will leave it at that and turn from transderivational economy constraints to the derivational economy constraint Last Resort.

3.4. Last Resort

3.4.1. The Constraint

Thus far, we have seen that if certain features are present on a lexical item, movement must apply so as to check them. As it stands, this leaves open the possibility that some instances of movement could also be triggered by some factors that are unrelated to feature checking, or, indeed, could occur without being triggered at all. This possibility is excluded by Last Resort. According to Last Resort, movement may apply only if it is triggered by morphological features. Thus, just like Fewest Steps, Last Resort prohibits truly optional, unforced movement. For the time being, I will assume the following version of Last Resort, which is based on Chomsky (1995, 280) (for predecessors, cf. Chomsky (1993, 33), Collins (1994a), Lasnik (1993, 12f; 1995, 620)), and Wilder & Čavar (1994, 60), among many others).

(42) *Last Resort*:
α is raised to a position β only if β is in the minimal residue of γ, and α enters into a checking relation with γ.

Thus, according to Last Resort, movement can apply only if it results in feature checking; given Procrastinate, it follows that overt movement (pre-Spell-Out) can take place only if either a feature of the moved item or a feature of the target head is strong and is checked by this movement. The *minimal residue* or *checking domain*, i.e., the domain in which a feature of the moved item can be checked with the target head, can be defined as follows (cf. Chomsky (1993, 11ff)).[22]

(43) *Minimal Residue* ('Checking Domain'):
The minimal residue of a head X is the smallest set M of categories α such that:
a. X-Adj, SpecX, and XP-Adj is in M.
b. If α is in M, then α-Adj is in M.

Thus, the minimal residue of some head comprises adjuncts to that head, its specifier, adjuncts to its maximal projection, anything adjoined to one of these positions, and so on.

In contrast to Procrastinate and Fewest Steps, Last Resort is not a global, transderivational principle. Rather, it is a local, derivational principle – at each step in the derivation, one can determine whether or not Last Resort is satisfied. In what follows, I will discuss some consequences of Last Resort for the movement types *wh*-movement, topicalization, and scrambling, and after that I will address the interaction between Last Resort and the Fewest Steps condition.

3.4.2. Wh-Movement, Topicalization, and Scrambling

Given Last Resort as defined in (42), overt *wh*-movement, topicalization, and scrambling must be triggered by strong features ([+wh], [top], and [scr], respectively). Following Chomsky (1995), I assume that the class of features that triggers movement is further divided into (a) features that are semantically interpretable (henceforth [+Interpretable]), like, e.g., the features [+N] and [+Plural] on a noun, and (b) features that are not semantically interpretable ([–Interpretable)], like, e.g., the feature [Case] on a noun. A crucial difference between these two classes of features is that only [–Interpretable] features are deleted by checking; [+Interpretable] features persist after checking and are accessible for further operations. .

Given that the features [top] and [scr] cannot be held responsible for a unique semantic effect (recall that this does not imply that topicalization and scrambling structures cannot differ from the respective in situ versions with respect to semantic interpretation), we can conclude that [top] and [scr] are both [–Interpretable]. In addition, since the concepts of LF topicalization or LF scrambling appear poorly motivated, these features are best assumed to be either strong on both the target head (Top, V or I) and the item that undergoes movement, or not present at all. Finally, the presence of [top] and [scr] on the target and on the moved item can be assumed to be optional throughout.

---

[22] This recursive definition is actually not the one given in Chomsky (1993), but it is equivalent to it.

Given that topicalization can only occur once per clause in languages like English and German, there can be only one [top] feature per Top head, and if the number of [top] features on lexical items does not match the number of [top] features on Top nodes in a given numeration, the derivation will invariably crash. The question arises of whether (and if so, how) [top] feature insertion on Top is related to V/2 movement to Top. Chomsky (1993, 32) suggests that there might indeed be a correlation that should be captured, so that [top] feature insertion becomes obligatory in V/2 contexts ("Raising of I to C may automatically make the relevant feature of C strong (the V/2 phenomenon)"). This view might be adequate for languages like German that do indeed show a tight relation of V/2 movement and topicalization;[23] however, it is arguably less adequate for topicalization in English (that generally occurs without V/2 movement, unless a negative operator is preposed), and for the derivation of VSO structures in the Celtic languages (assuming that they involve I-to-C (I-to-Top) movement). I will have to leave this issue open; cf. Watanabe (1993) for extensive discussion.

In contrast to topicalization, scrambling can apply multiply in a single checking domain, and given that [scr] is [–Interpretable], this implies that there is no restriction on the number of [scr] features that a target node (V or I in German) can bear.[24]

Turning next to *wh*-movement, I will make two assumptions, both of which deviate from what is proposed in Chomsky (1995). First, [+wh] features, in contrast to [top] or [scr], are not inherently [–Interpretable] or [+Interpretable]. Rather, the [+wh] feature of the target C is [+Interpretable], whereas the [+wh] feature of *wh*-phrases is [–Interpretable].[25] Second, in languages like English and German, the [+wh] fea-

---

[23]The question arises of whether subject-initial V/2 clauses in German also involve [top] feature checking in SpecTop, with the finite verb in Top (which essentially corresponds to the standard view; cf. Gärtner & Steinbach (1994) and Schwartz & Vikner (1996)), or whether subject and finite verb are in SpecI and (a left-peripheral) I, respectively (as argued for by Travis (1984; 1991) and Zwart (1993)). Although this question is strictly speaking orthogonal to most of the issues that concern me in this study, I will adopt the former view here. This view is reinforced by the fact that *wh*-subjects behave like *wh*-objects in that they cannot be topicalized in German. Under the approach developed in subsection 5.2.1 below, this implies that *wh*-subjects have participated in [top] feature checking in SpecTop.

[24]An empirically adequate way to account for the fact that English does not exhibit scrambling would simply be to state that the insertion of [scr] on any target node is blocked in this language. Similarly, the fact that Russian seems to have a larger set of landing sites for scrambling than German – it also permits scrambling to NP and CP (cf. Müller (1995, ch. 3)) – can be accounted for by assuming that N and C can check [scr] features, in addition to V and I (more generally, the heads of the extended V projection). Of course, this, by itself, does not yet explain cross-linguistic variability with respect to scrambling options; it is merely a way of implementing whatever explanation ultimately proves correct in the present framework. See the literature cited in footnote 9 of chapter 1 for some relevant proposals.

[25]Evidence for this classification of the [+wh] features of a *wh*-phrase and of C is provided by semantic considerations. Based on Karttunen's (1977) semantics of questions, Heim (1992), Beck (1996), and Stechow (1996) argue for LF representations where the *wh*-phrase is viewed as an ordinary, existentially quantified XP (i.e., the [+wh] feature of the *wh*-phrase is ignored for interpretation), whereas the [+wh] feature of C systematically contributes to interpretation – it introduces what Heim (1992) calls the "interro-

tures of *wh*-phrases are always weak. The [+wh] feature that triggers movement here is that on C, even though this feature is [+Interpretable]. Of course, given that the [+wh] feature of C is initially both [+Interpretable] and strong in the languages under discussion, we have to make the auxiliary assumption that after checking, this feature becomes weak; otherwise, PF convergence could never be attained. Thus, if the [+wh] feature of a *wh*-phrase is checked by movement, it disappears (being [–Interpretable]). But if the [+wh] feature of C is checked in an English-type language, it does not disappear (being [+Interpretable]). Rather, it remains in the derivation and is accessible for further checking of [+wh] features on *wh*-phrases, but, having turned weak, only via LF *wh*-movement (given Procrastinate).

From these assumptions, it follows that only one *wh*-phrase can undergo overt movement to SpecC in a multiple question in languages like English and German; all remaining *wh*-phrases undergo raising to the checking domain of C at LF:[26]

(44) a. (I wonder) $[_{CP}$ – C$_{[}$+wh] $[_{IP}$ John gave what$_{[}$+wh] to whom$_{[}$+wh]] (D-structure)
b. (I wonder) $[_{CP}$ what C$_{[}$+wh] John gave t to whom$_{[}$+wh]] (S-structure)
c. (I wonder) $[_{CP}$ what to whom C$_{[}$+wh] $[_{IP}$ John gave t t]] (LF)

Next, consider an example that is ambiguous between a reading as a multiple matrix question and a multiple embedded question (see Baker (1970)):

(45) Who$_1$ C t$_1$ wonders $[_{CP}$ where$_2$ C we bought what$_3$ t$_2$ $]_?$

In this sentence, there are three *wh*-phrases, two of which have undergone overt movement, and one of which remains in situ at S-structure (raising to either the embedded or the matrix SpecC position at LF, cf. Lasnik & Saito (1984; 1992), among others). Both C heads bear a strong [+wh] feature in the numeration. These strong [+wh] features are checked by the two instances of overt *wh*-movement; as a result, the [+wh] features on the C nodes turn weak, but (being [+Interpretable]) remain accessible for further checking at LF. Subsequent LF raising of the in situ *wh*-phrase *what*$_3$ is then free to target either the matrix or the embedded C node, thereby producing the two possible readings.

Finally, it remains to be shown how language-specific variation with respect to *wh*-movement (the "Wh-Movement Parameter") can be accounted for. Assuming both the [+wh] feature of C and the [+wh] features of *wh*-phrases to be weak produces the

---

gativizer," and it is ultimately responsible for the fact that questions are interpreted as sets of propositions (and not as propositions, as declarative clauses are).

[26] At this point, the advantage of locating strength of [+wh] in C in languages of the English type, rather than in the *wh*-phrase, becomes apparent. If overt *wh*-movement were triggered by a strong feature on the *wh*-phrase, we would have to ensure by resort to an ad hoc assumption to the effect that exactly *n* *wh*-phrases bear a strong [+wh] feature if there are *n* [+wh] C heads in a numeration; cf. Lasnik (1995, 621).

right consequences for *wh*-in situ languages of the Korean/Japanese type; here, all *wh*-movement is postponed to the covert component. In languages that exhibit multiple S-structure *wh*-movement to the checking domain of C (like Bulgarian; see Rudin (1988)), we can assume that the [+wh] features of *wh*-phrases are inherently strong; it will then be irrelevant whether the [+wh] feature of C is (initially) strong, as in English, or weak, as in Korean and Japanese.

To sum up, the following system has emerged for German: The features [top] and [scr] are [–Interpretable] and strong on both the target head and the item that undergoes movement. The feature [+wh] is [+Interpretable] and strong on the target head C and [–Interpretable] and weak on the *wh*-phrase.

3.4.3. The Interaction of Last Resort and Fewest Steps

The derivational Last Resort condition (repeated here in (46)) and the transderivational Fewest Steps condition (repeated in (47)) are formally very different.

(46) *Last Resort*:
α is raised to a position β only if β is in the minimal residue of γ, and α enters into a checking relation with γ.

(47) *Fewest Steps*:
If two derivations D_1 and D_2 are in the same reference set and D_1 involves fewer operations than D_2, then D_1 is to be preferred over D_2.

Nevertheless, they bear a certain resemblance to each other in that they both prohibit truly optional, unforced movement. Therefore, the question arises of whether one of the two constraints could be dispensed with. Now, a well-known fact about global, transderivational constraints is that they raise the problem of computational complexity – comparing a very large (possibly infinite) number of derivations is not an easy task. Hence, global constraints like Fewest Steps are arguably inferior to local constraints like Last Resort from a conceptual point of view.[27] Given this state of affairs, we have to ask whether the effects of Fewest Steps can be derived from the Last Resort condition. In what follows, I would like to show that this is not the case.

Consider first the case of illicit *wh*-topicalization in English, discussed in subsection 3.2.1 above. The relevant example is repeated here in (48) (where (48-a) is the S-structure representation, and (48-b) is the representation of the sentence at LF):

(48) a. *Who$_1$ t$_1$ said [$_{CP}$ that [$_{TopP}$ who$_2$ Top [$_{IP}$ John likes t$_2$]]] ? (S-structure)
b. [$_{SpecC}$ [$_{SpecC}$ who$_1$] who$_2$] t$_1$ said [$_{CP}$ that [$_{IP}$ John likes t$_2$]] ? (LF)

[27] See especially Collins (1997) and Frampton & Gutmann (1997), but also Chomsky (1995) for some pertinent remarks.

As shown above, this derivation is ruled out by Fewest Steps (because there is another derivation in the same reference set which does not involve overt topicalization of the *wh*-phrase *who*$_2$). But Last Resort is not violated in this derivation: Suppose that the *wh*-phrase *who*$_2$ has both a (strong) [top] feature and a (weak) [+wh] feature, and that there is a strong [top] feature on the embedded Top node and a strong [+wh] feature on the matrix C node (that becomes weak after checking with the *wh*-phrase *who*$_1$). Under these assumptions, the *wh*-phrase *who*$_2$ may perfectly well undergo topicalization before Spell-out and subsequent LF raising to the matrix SpecC position, as Last Resort requires – the overt movement step of the *wh*-phrase is licensed by its entering into a checking relation with the embedded Top node and the covert movement step by its entering into a checking relation with the matrix C node. Thus, it seems that Last Resort could rule out a derivation like (48) only under an assumption like (49):[28]

(49) The features [$_{\text{top}]}$ and [+wh] cannot co-occur on a lexical item.

But it would clearly be preferable to derive such a condition rather than just state it as such. The Fewest Steps condition does just that – in contrast to Last Resort, it explains why (49) seems to be descriptively correct. An identical reasoning applies in the case of illicit *wh*-scrambling; without Fewest Steps, we would have to resort to an explicit statement to the effect that the features [scr] and [+wh] cannot co-occur on a lexical item.

As a second example illustrating that the effects of Fewest Steps cannot be derived by Last Resort, consider again the case of argument/adjunct asymmetries in *wh*-movement, as discussed in subsection 3.2.5. A relevant sentence is repeated in (50):

(50) *How$_3$ do you wonder [$_{\text{CP}_1}$ whether John said [$_{\text{CP}_2}$ t_3' C Mary solved the problem t_3]] ?

In order to ensure that (50) is strongly (ECP-like) ungrammatical, LF deletion of the "offending" trace t_3' must be precluded. As noted above, such a deletion operation is in fact ruled out by Fewest Steps (because the chain $<$*how*$_3$,t_3',$t_3$$>$ is already a legitimate LF object), but Last Resort does not say anything about this case.

Summarizing, it turns out that the transderivational economy constraint Fewest Steps cannot be reduced to the derivational economy constraint Last Resort, notwithstanding the general conceptual problem that might be involved in the postulation of global constraints.

[28](49) is a feature co-occurrence restriction of exactly the type that played an important role in work carried out within the GPSG framework; see, e.g., Gazdar et al. (1985, 27ff). I will later (in chapter 6) return to restrictions of this kind.

4. Cyclicity Constraints

So far, I have discussed locality constraints and economy constraints. Now I will turn to a third type of constraint, viz., cyclicity constraints. Under this label I will subsume constraints that say something about legitimate and illegitimate interactions of movement operations. In particular, I will be concerned with two central notions, viz. (a) *strict cyclicity* and (b) *successive cyclicity*. The constraints corresponding to these notions will be the *Strict Cycle Condition* and the *Principle of Unambiguous Binding* (PUB). I will turn first to strict cyclicity and will address successive cyclicity and the related concept of improper movement later.

4.1. Strict Cyclicity

The notion of strict cyclicity will be important in the remainder of this monograph and will be shown to play a crucial role in deriving certain peculiarities of remnant movement. For this reason, I would like to examine this concept in somewhat more detail. I will start with a brief history of the Strict Cycle Condition, beginning with the situation in the early seventies, i.e., in a framework in which traces had not yet been discovered.[29]

4.1.1. The Strict Cycle Condition before Trace Theory

As a point of departure, consider again typical violations of the *Wh*-Island Condition, as discussed in subsections 2.1 and 3.2.5 above.[30]

(51) a. *How$_1$ do [$_{IP}$ you wonder [$_{CP}$ which car$_2$ [$_{IP}$ to fix t$_2$ t$_1$]]] ?
b. ??Which car$_2$ do [$_{IP}$ you wonder [$_{CP}$ how$_1$ [$_{IP}$ to fix t$_2$ t$_1$]]] ?

According to standard reasoning (cf. Chomsky (1973; 1977)), (51-a) and (51-b) violate the Subjacency Condition that blocks a crossing of more than one bounding node: The embedded SpecC position is occupied, and therefore *how* in (51-a) and *which car* in (51-b) must undergo long-distance movement in one fell swoop. This movement step crosses two IPs. IP is a bounding node (in English), and Subjacency is violated.

However, the following derivation of a sentence like (51-b) is not ruled out by Subjacency:

---

[29] The discussion will be confined to the interaction of the Strict Cycle Condition and locality. See Pullum (1979; 1992) for a comprehensive critical discussion of (strict) cyclicity in syntax that also touches many other relevant issues, and for some speculations as to the conceptual foundations of the notion of cyclicity, which also seems to play a role in non-linguistic systems. Considerations along these lines may ultimately prove to be illuminating, but they are outside the scope of the present study.

[30] I have inserted traces in the initial positions here, but this is only for convenience, in order to indicate that *how* and *which car* originate in the embedded clause.

(52) a. *D-structure*:
[$_{CP}$ – do [$_{IP}$ you wonder [$_{CP}$ – [$_{IP}$ PRO to fix which car$_2$ how$_1$]]]] ?
b. *Wh-Movement of 'which car' to the embedded SpecC position*:
[$_{CP}$ – do [$_{IP}$ you wonder [$_{CP}$ which car$_2$ [$_{IP}$ PRO to fix how$_1$]]]] ?
c. *Wh-Movement of 'which car' to the matrix SpecC position*:
[$_{CP}$ which car$_2$ do [$_{IP}$ you wonder [$_{CP}$ [$_{IP}$ PRO to fix how$_1$]]]] ?
d. *Wh-Movement of 'how' to the embedded SpecC position*:
[$_{CP}$ which car$_2$ do [$_{IP}$ you wonder [$_{CP}$ how$_1$ [$_{IP}$ PRO to fix]]]] ?

The reason is that at no point in the derivation are two bounding nodes (IPs) crossed; each movement operation crosses only one IP. In a nutshell, the Subjacency condition is rendered vacuous in this derivation because the *wh*-island erected by the *wh*-phrase in the embedded SpecC position is created only *after* long-distance movement has applied. What, then, excludes the derivation in (52)?

In order to rule out a derivation of examples involving a *wh*-island effect as in (52), Chomsky (1973, 243) introduces the Strict Cycle Condition, which is given in (53) in its original form:

(53) *Strict Cycle Condition*:
No rule can apply to a domain dominated by a cyclic node α in such a way as to affect solely a proper subdomain of α dominated by a node β which is also a cyclic node.

Given (53), it is necessary to define what cyclic nodes are. For now, it may suffice to assume, basically following Chomsky (1973), that CP and NP are cyclic nodes in English. With respect to the derivation in (52), the Strict Cycle Condition then seems to make the correct prediction: *Wh*-movement of *how* in the last step applies to the domain dominated by the matrix CP (a cyclic node), but since it ends up in a position dominated by the embedded CP (which is a cyclic node too), it affects only a proper subdomain of the matrix CP dominated by the embedded CP – and this is prohibited by the Strict Cycle Condition.

Still, there is more to be said. Consider the following sentence, where *wh*-movement takes place in an embedded CP.

(54) [$_{CP_1}$ I know [$_{CP_2}$ what$_3$ John did t$_3$]]

Again, the matrix CP and the embedded CP are cyclic nodes. However, with exactly the same reasoning as in the case of the illicit derivation in (52), it seems that we have to conclude that *wh*-movement to the embedded SpecC position in (54) also violates the Strict Cycle Condition. To avoid this unwelcome conclusion, we must clarify under which conditions a rule "applies to a domain dominated by a cyclic node" in the sense of the Strict Cycle Condition (53). More generally, it seems that the

Strict Cycle Condition must be amended by adding a notion of the *cycle*.[31] Basically, the cycle says that movement operations apply from bottom to top. Ignoring NPs, this means that in a structure of the form in (55), all movement applies first to the most deeply embedded cyclic node (the CP_3 cycle), next to the CP_2 cycle, and finally to the CP_1 cycle.

(55) $[_{\mathrm{CP}_1} \ldots [_{\mathrm{CP}_2} \ldots [_{\mathrm{CP}_3} \ldots] \ldots] \ldots]$

With this in mind, let us consider again the problematic example in (54) and try to determine whether or not the Strict Cycle Condition still rules out this example. Given that rule application applies cyclically, from bottom to top, *wh*-movement in (54) may apply on the CP_2 cycle, i.e., at a point in the derivation at which the CP_1 cycle has not yet been reached. Hence, *wh*-movement of *what*$_3$ in (54) does not "apply to a domain dominated by a cyclic node," although the matrix CP_1 *is* in principle a cyclic node.

To illustrate this, consider a sentence that differs minimally from (54) in that the matrix subject is not in situ, but itself undergoes *wh*-movement:[32]

(56) $[_{\mathrm{CP}_1}$ Who knows $[_{\mathrm{CP}_2}$ what John did $]]_{?}$

Given that there are two instances of *wh*-movement in (56), there are two derivations that are possible a priori. In derivation D_1, movement of *who* precedes movement of *what*, and in derivation D_2, movement of *what* precedes movement of *who*. Derivation D_1 is given in (57):

(57) a. *D-structure*:
$[_{\mathrm{CP}_1} - [_{\mathrm{IP}}$ who$_4$ knows $[_{\mathrm{CP}_2} - [_{\mathrm{IP}}$ John did what$_3$]]]]
b. *Wh-Movement applies to 'who' – CP_1 cycle*:
$[_{\mathrm{CP}_1}$ who$_4$ $[_{\mathrm{IP}}$ t$_4$ knows $[_{\mathrm{CP}_2} - [_{\mathrm{IP}}$ John did what$_3$]]]]
c. *Wh-Movement applies to 'what' – CP_1 cycle*:
$[_{\mathrm{CP}_1}$ who$_4$ $[_{\mathrm{IP}}$ t$_4$ knows $[_{\mathrm{CP}_2}$ what$_3$ $[_{\mathrm{IP}}$ John did t$_3$]]]]

Since *wh*-movement in the matrix clause CP_1 implies that all subsequent movement operations occur on the CP_1 cycle, subsequent *wh*-movement in the embedded clause violates the Strict Cycle Condition. Thus, D_1 in (57) is not a possible derivation of (56). However, D_2 in (58) is:

(58) a. *D-structure*:
$[_{\mathrm{CP}_1} - [_{\mathrm{IP}}$ who$_4$ knows $[_{\mathrm{CP}_2} - [_{\mathrm{IP}}$ John did what$_3$]]]

[31] See Perlmutter & Soames (1979), where this problem is discussed in some detail.

[32] The same point could be made on the basis of (54) under the Predicate-Internal Subject Hypothesis, i.e., under the assumption that the subject NP *I* in (54) is base-generated VP-internally and then raised to SpecI.

b. *Wh-Movement applies to 'what' – CP_2 cycle*:
[$_{CP_1}$ – [$_{IP}$ who$_4$ knows [$_{CP_2}$ what$_3$ [$_{IP}$ John did t$_3$]]]
c. *Wh-Movement applies to 'who' – CP_1 cycle*:
[$_{CP_1}$ who$_4$ [$_{IP}$ t$_4$ knows [$_{CP_2}$ what$_3$ [$_{IP}$ John did t$_3$]]]

Here, *wh*-movement in the embedded clause applies first. Since it does not affect a position in the matrix clause, we can assume that this movement still applies exclusively on the CP_2 cycle, in accordance with the Strict Cycle Condition. Subsequent *wh*-movement in the matrix clause then applies on the CP_1 cycle but obviously does not violate the Strict Cycle Condition.

Thus, given these assumptions, the Strict Cycle Condition in (53) can be reformulated in a slightly more transparent way as in (59) (cf. Freidin (1992, 107)):

(59) *Strict Cycle Condition* (revised version):
No rule may apply solely within a cyclic subdomain of the current cycle.

With this in mind, let us return to (52), which gives the illicit derivation of a typical example involving a violation of the *Wh*-Island Condition. This derivation is repeated in (60):

(60) a. *D-structure*:
[$_{CP}$ – do [$_{IP}$ you wonder [$_{CP}$ – [$_{IP}$ PRO to fix which car$_2$ how$_1$]]]] ?
b. *Wh-Movement of 'which car' to the embedded SpecC position*:
[$_{CP}$ – do [$_{IP}$ you wonder [$_{CP}$ which car$_2$ [$_{IP}$ PRO to fix how$_1$]]]] ?
c. *Wh-Movement of 'which car' to the matrix SpecC position*:
[$_{CP}$ which car$_2$ do [$_{IP}$ you wonder [$_{CP}$ [$_{IP}$ PRO to fix how$_1$]]]] ?
d. *Wh-Movement of 'how' to the embedded SpecC position*:
[$_{CP}$ which car$_2$ do [$_{IP}$ you wonder [$_{CP}$ how$_1$ [$_{IP}$ PRO to fix]]]] ?

Note that the second movement operation (*wh*-movement of *which car* to the matrix SpecC position) applies on the matrix CP cycle. Given the notion of the cycle, all subsequent movement operations apply on this very same cycle. But then, it follows that *wh*-movement of *how* to the embedded SpecC position in the last step applies on the matrix cycle, but affects only a proper cyclic subdomain, viz., the embedded cycle, in violation of the Strict Cycle Condition.

4.1.2. The Strict Cycle Condition and Trace Theory

Thus far, I have shown that the Strict Cycle Condition appears to be necessary, in interaction with a locality constraint like Subjacency, in order to rule out examples which exhibit a *Wh*-Island Condition violation. However, with the emergence of trace theory and S-structure filters, this conclusion does not appear to be valid anymore. Thus, Freidin (1978) shows that given (a) trace theory and/or (b) a representational view of Subjacency, the Strict Cycle Condition is superfluous, as far as the derivation of *wh*-island effects is concerned. Let me first reconsider the derivation in (52) under the assumption that all movement leaves traces. (52) then looks as in (61):

(61) a. *D-structure*:
$[_{CP}$ – do $[_{IP}$ you wonder $[_{CP}$ – $[_{IP}$ PRO to fix which car$_2$ how$_1$]]]] ?
b. *Wh-Movement of 'which car' to the embedded SpecC position*:
$[_{CP}$ – do $[_{IP}$ you wonder $[_{CP}$ which car$_2$ $[_{IP}$ PRO to fix t$_2$ how$_1$]]]] ?
c. *Wh-Movement of 'which car' to the matrix SpecC position*:
$[_{CP}$ which car$_2$ do $[_{IP}$ you wonder $[_{CP}$ t'_2 $[_{IP}$ PRO to fix t$_2$ how$_1$]]]] ?
d. *Wh-Movement of 'how' to the embedded SpecC position is impossible*:
$[_{CP}$ which car$_2$ do $[_{IP}$ you wonder $[_{CP}$ (t'_2) (how$_1$) $[_{IP}$ PRO to fix t$_2$ t$_1$]]]] ?

As indicated, *wh*-movement of *how* to the embedded SpecC position in the last step is not possible anymore; the reason is that this position is already occupied by an intermediate trace of the preceding movement operation which applied to *which car*. Hence, an appeal to the Strict Cycle Condition is no longer necessary in order to rule out the derivation in (61).

Next, suppose that movement does not (have to) leave traces in intermediate positions or that such traces can freely be overwritten by later operations. Suppose, however, that the Subjacency condition is formulated not derivationally, but rather as an S-structure filter. Reconsider the derivation in (52) with this assumption in mind:

(62) a. *D-structure*:
$[_{CP}$ – do $[_{IP}$ you wonder $[_{CP}$ – $[_{IP}$ PRO to fix which car$_2$ how$_1$]]]] ?
b. *Wh-Movement of 'which car' to the embedded SpecC position*:
$[_{CP}$ – do $[_{IP}$ you wonder $[_{CP}$ which car$_2$ $[_{IP}$ PRO to fix t$_2$ how$_1$]]]] ?
c. *Wh-Movement of 'which car' to the matrix SpecC position*:
$[_{CP}$ which car$_2$ do $[_{IP}$ you wonder $[_{CP}$ (t'_2) $[_{IP}$ PRO to fix t$_2$ how$_1$]]]] ?
d. *Wh-Movement of 'how' to the embedded SpecC position*:
$[_{CP}$ which car$_2$ do $[_{IP}$ you wonder $[_{CP}$ how$_1$ $[_{IP}$ PRO to fix t$_2$ t$_1$]]]] ?

If Subjacency applies at S-structure, i.e., after all overt movement operations have taken place, it can apply only to the output of the last operation in (62). But, clearly, the resulting representation violates Subjacency, with *which car* being separated from its trace t_2 by two bounding nodes (IPs). Thus, again, an appeal to the Strict Cycle Condition appears to be unnecessary. Summarizing, assuming that all movement must leave a trace (that cannot be deleted subsequently), *or* that the Subjacency condition applies at S-structure, makes the Strict Cycle Condition superfluous (at least as far as *wh*-island effects are concerned).

4.1.3. The Strict Cycle Condition and Principles-&-Parameters Theory

Interestingly, both these assumptions are dropped in the versions of movement theory that were later developed by Lasnik & Saito (1984; 1992) and Chomsky (1986), among many others. In these approaches, it is postulated (a) that movement (of arguments) does not have to leave intermediate traces (and that, moreover, intermediate argument traces can be deleted systematically) and (b) that Subjacency is a derivational constraint applying between D-structure and S-structure. Evidently, according to this view, the Strict Cycle Condition is indispensable for an account of the *Wh*-Island Condition, at least in cases where the *wh*-phrase undergoing long-distance movement is an argument.

4.1.4. *The Strict Cycle Condition and the Minimalist Program*

Finally, let me turn to the fate of the Strict Cycle Condition in the minimalist program. As shown in (63), it emerges that the Strict Cycle Condition is necessary in the minimalist program in order to derive the effects of the *Wh*-Island Condition.

(63) a. *D-structure*:
[$_{CP}$ – do [$_{IP}$ you wonder [$_{CP}$ – [$_{IP}$ PRO to fix which car$_2$ how$_1$]]]] ?
b. *Form Chain applies long-distance to 'which car'*:
[$_{CP}$ which car$_2$ do [$_{IP}$ you wonder [$_{CP}$ t'_2 [$_{IP}$ PRO to fix t_2 how$_1$]]]] ?
c. *Form Chain applies to 'how'*:
[$_{CP}$ which car$_2$ do [$_{IP}$ you wonder [$_{CP}$ how$_1$ [$_{IP}$ PRO to fix t_2 t_1]]]] ?

The step in (63-c) must be ruled out by the Strict Cycle Condition because (a) Subjacency cannot exist as an S-structure filter (there is no S-structure level anymore in the minimalist program) and (b) embedded Form Chain in (63-c) can overwrite and thereby delete the intermediate argument trace. Note that this trace t'_2 must be deleted anyway, since otherwise an illegitimate LF-object (a non-uniform chain, cf. subsection 3.2.5) would result (Chomsky & Lasnik (1993)).

Furthermore, it turns out that unlike what is the case in Principles-&-Parameters Theory, the Strict Cycle Condition is necessary even for cases of adjunct movement across a *wh*-island, as in (51-a), which is repeated here as (64).

(64) *How$_1$ do [$_{IP}$ you wonder [$_{CP}$ which car$_2$ [$_{IP}$ to fix t_2 t_1]]] ?

The relevant derivation is given in (65).

(65) a. *D-structure*:
[$_{CP}$ – do [$_{IP}$ you wonder [$_{CP}$ – [$_{IP}$ to fix which car$_2$ how$_1$]]]] ?
b. *Form Chain applies long-distance to 'how'*:
[$_{CP}$ how$_1$ do [$_{IP}$ you wonder [$_{CP}$ t'_1 [$_{IP}$ to fix which car$_2$ t_1]]]] ?
c. *Form Chain applies to 'which car'*:
[$_{CP}$ how$_1$ do [$_{IP}$ you wonder [$_{CP}$ which car$_2$ [$_{IP}$ to fix t_2 t_1]]]] ?

Again, without the Strict Cycle Condition, nothing prevents *which car* from moving to the embedded SpecC position in the last step. The moved *wh*-phrase overwrites and thereby deletes the intermediate adjunct trace t'_1. This is not possible under standard assumptions in Principles-&-Parameters Theory (see footnote 17), but it is under the guiding assumptions of the minimalist program. Recall that the deletion of intermediate adjunct traces is not blocked by stipulations concerning the level of γ-marking anymore, but rather ruled out by the Fewest Steps condition (cf. subsection 3.2.5 above). The Fewest Steps condition, however, does not preclude the operation in (65-b), where adjunct trace deletion comes at no cost as a result of a single operation Form Chain applying to the argument *wh*-phrase *which car*$_2$.

Thus, given that something like the Strict Cycle Condition seems to be called for in the minimalist program, the question arises of whether its effects can be derived from independent, more basic assumptions. The question is answered in the affirmative in Chomsky (1993); in the following subsection, I summarize that proposal.

4.1.5. Deriving the Strict Cycle Condition from the Extension Condition

Chomsky (1993, 22ff) assumes that syntactic trees are built by successive applications of generalized transformations; this operation is called "Merge" in Chomsky (1995) (where some differing properties are attributed to it, which do not concern me here). Movement is viewed as an instance of Merge. The structure-building operation Merge can be defined as follows:

(66) *Generalized Transformation* (Merge):
Merge is a binary operation. Given two trees α, β, Merge targets β, adds an empty element ϕ, and replaces ϕ by α, thereby creating a new tree γ, with β as a proper subpart.

On the basis of (66), movement (Move α) can be understood as Merge with the additional requirement that α is a proper subpart of β.

In addition to that, Chomsky (1993) assumes that Merge (hence, movement) is further restricted by an *Extension Condition*, as in (67).

(67) *Extension Condition for Merge*:
ϕ in (66) must be external, unless
a. the result of Merge is an adjunction structure (head movement, XP adjunction); or
b. Merge is a movement operation after Spell-Out (LF movement)

From (67) it follows that substitution operations have to extend the phrase marker (this, of course, presupposes that movement and pure structure-building operations may alternate, cf. footnote 1 above). This derives the effect of the Strict Cycle Condition (53) (or (59)), as far as the *Wh*-Island Condition is concerned. To see this, consider again the relevant step of an acyclic derivation of a sentence exhibiting a violation of the *Wh*-Island Condition, as in (63), repeated here.

(68) a. [$_{CP}$ which car$_2$ do [$_{IP}$ you wonder [$_{CP}$ t$'_2$ [$_{IP}$ PRO to fix t$_2$ how$_1$]]]] ?
b. [$_{CP}$ which car$_2$ do [$_{IP}$ you wonder [$_{CP}$ how$_1$ [$_{IP}$ PRO to fix t$_2$ t$_1$]]]] ?

The movement operation in (68-b) which puts the *wh*-adjunct *how* into the embedded SpecC position, overwriting the intermediate trace of *which car*, does not meet the requirement imposed by the Extension Condition on Merge. The reason is that embedded *wh*-movement of *how* in (68-b) neither is an LF operation nor creates an adjunction structure – it is S-structure substitution in SpecC, and as such illegitimately fails to extend the tree built up so far.

Thus, the Extension Condition on generalized transformations (therefore, on movement) put forward in Chomsky (1993) at first sight seems to derive the effects of the Strict Cycle Condition. However, note that it rests on two assumptions that will turn out to be problematic – first, head movement (and XP adjunction) must be exempted from the Extension requirement; and second, LF movement is an exception to this condition on Merge. In the following subsection, I will show why these two exceptions must be made, and I will argue that they are empirically problematic.

4.1.6. Against Deriving the Strict Cycle Condition from the Extension Condition

Head Movement The two exceptions in the Extension Condition (67) with respect to adjunction operations and LF operations are unavoidable. This is quite obvious for the case of head movement (analyzed as adjunction to some X^0 category): It is impossible that adjunction of one head to some other head extends the tree built up so far. This is shown in (69).

(69) a. $[_{\beta'}\ \beta$ [... α ...]]
b. $[_{\beta'}\ [_{\beta}\ \alpha\ \beta$] [... t_α ...]]

Here, (69-a) is the structure created by Merge prior to head movement; the head α must be included in the complement of the target head β.[33] But, clearly, if α is included in the complement of β in the pre-movement structure, this implies that the minimal projection of β, viz., β', is also present before head movement applies. Therefore, head movement as in (69-b) invariably fails to meet the Extension requirement on Merge – the landing site of the moved head α is always below the minimal projection of β. Hence, an exception to the Extension Condition (67) like the one accomplished by clause (a) appears to be crucial.[34]

There is, however, conflicting evidence. As noted by Marantz (1995, 362) and Baker (1988, ch. 7)), there is good reason to assume that head movement does indeed obey strict cyclicity. If this is the case, we have to conclude that the Strict Cycle Condition applies to head movement after all and that the idea to derive strict cyclicity from the Extension Condition is not on the right track.

---

[33]This is the typical structure in which head movement is legitimate, see Travis (1984), Chomsky (1986), Baker (1988).

[34]As noted by Bobaljik & Brown (1997), this exception to the Extension Condition can be dispensed with if one is willing to make the assumption that head movement of an item α that is included in a tree β may target a distinct tree γ that is unrelated to β at the time when movement of α applies (an "interarboreal" operation). I will not adopt this assumption here.

LF Movement LF movement usually does not extend the phrase marker. This is evident for Case-driven LF movement of an object NP in English (assuming for the purpose of the discussion that there is such a thing) and also for *wh*-movement at LF. I will show this on the basis of the *wh*-island effect with LF movement that occurs in *wh*-in situ languages like Chinese, Japanese, and Korean (cf. Huang (1982), Lasnik & Saito (1984; 1992, 30ff), and Koster (1987, 216ff), among many others). For concreteness, consider the following example from Korean (Shin-Sook Kim, p.c.):

(70) [$_{\text{CP}}$ – nŏ-nŭn [$_{\text{CP}}$ – Suna-ka muŏs-ŭl$_2$ wae$_1$ sa-ss-nŭnchi] alko-sip-ni] ?
you$_{top}$ Suna$_{nom}$ what$_{acc}$ why buy-PAST-Q want-to-know-Q

Here, the matrix C node and the embedded C node bear [+wh] features (as evidenced by the presence of the Q-morpheme in both clauses), and there are two *wh*-phrases in situ in the embedded clause. One option is to interpret (70) as a matrix yes/no question with an embedded multiple question. This reading corresponds to an LF representation in which both *wh*-phrases undergo local movement within the embedded clause at LF, and it is not relevant in the present context. The two interesting LF representations to be considered here involve LF movement of one *wh*-phrase to the embedded SpecC position and LF movement of the other *wh*-phrase to the matrix SpecC position. As shown in (71), only the LF representation in which the *wh*-argument *muŏs-ŭl* ('what') has undergone long movement and the *wh*-adjunct *wae* ('why') has undergone short movement corresponds to an available reading.

(71) a. [$_{\text{CP}}$ muŏs-ŭl$_2$ nŏ-nŭn [$_{\text{CP}}$ wae$_1$ Suna-ka t$_2$ t$_1$ sa-ss-nŭnchi] alko-sip-ni] ?
'What do you want to know why Suna bought?'
b. *[$_{\text{CP}}$ wae$_1$ nŏ-nŭn [$_{\text{CP}}$ muŏs-ŭl$_2$ Suna-ka t$_2$ t$_1$ sa-ss-nŭnchi] alko-sip-ni] ?
'What is the reason such that you want to know what Suna bought for that reason?'

It seems that in order to rule out the LF representation in (71-b), two derivations must be excluded, as in the case of *wh*-island effects with movement before Spell-Out.[35]

Consider first the derivation in (72).

(72) a. *S-structure*:
[$_{\text{CP}}$ – nŏ-nŭn [$_{\text{CP}}$ – Suna-ka muŏs-ŭl$_2$ wae$_1$ sa-ss-nŭnchi] alko-sip-ni] ?
b. *Form Chain applies to 'muŏs-ŭl$_2$'*:
[$_{\text{CP}}$ – nŏ-nŭn [$_{\text{CP}}$ muŏs-ŭl$_2$ Suna-ka t$_2$ wae$_1$ sa-ss-nŭnchi] alko-sip-ni] ?
c. *Form Chain applies long-distance to 'wae$_1$'*:
[$_{\text{CP}}$ wae$_1$ nŏ-nŭn [$_{\text{CP}}$ muŏs-ŭl$_2$ Suna-ka t$_2$ t$_1$ sa-ss-nŭnchi] alko-sip-ni] ?

[35] What is said about the Korean example (70) in what follows carries over directly to analogous examples in Chinese and Japanese. Note in passing that the wellformedness of the LF representation in (71-a) is accounted for if we assume, following Huang (1982), Lasnik & Saito (1984; 1992), and Chomsky (1986), among others, that LF movement of arguments is immune to intervening *wh*-islands.

The last step of this derivation is in conflict with locality requirements, more precisely, with the Barriers Condition (see (1)). Next, consider a derivation where movement of *muŏs-ŭl* and movement of *wae* apply in reverse order:

(73) a. *S-structure*:
[$_{CP}$ – nŏ-nŭn [$_{CP}$ – Suna-ka muŏs-ŭl$_2$ wae$_1$ sa-ss-nŭnchi] alko-sip-ni] ?
b. *Form Chain applies long-distance to 'wae$_1$'*:
[$_{CP}$ wae$_1$ nŏ-nŭn [$_{CP}$ t'_1 Suna-ka muŏs-ŭl$_2$ t_1 sa-ss-nŭnchi] alko-sip-ni] ?
c. *Form Chain applies to 'muŏs-ŭl$_2$'*:
[$_{CP}$ wae$_1$ nŏ-nŭn [$_{CP}$ muŏs-ŭl$_2$ Suna-ka t_2 t_1 sa-ss-nŭnchi] alko-sip-ni] ?

This derivation does not violate locality constraints, since LF movement of *wae* applies successive-cyclically, via the embedded SpecC position. As before, the intermediate trace of the *wh*-adjunct t'_1 can be overwritten in the last step of the derivation by the *wh*-argument *muŏs-ŭl*$_2$. Given that LF movement does not have to meet the Extension Condition (67), it seems that this derivation does not violate any principle and that, therefore, (71-b) should be an admissible LF representation, contrary to fact. However, the last step in the derivation (73) can be excluded if LF movement obeys the Strict Cycle Condition, in analogy to what has been argued for above with regard to overt *wh*-movement across a *wh*-island.

For these reasons, I conclude that the Strict Cycle Condition is necessary for all kinds of movement (substitution and adjunction, movement before and after Spell-Out) in the minimalist program and that its effects cannot be derived from something like the Extension Condition on generalized transformations.[36]

4.1.7. Formulating the Strict Cycle Condition in the Minimalist Program

Given this state of affairs, it remains to formulate the Strict Cycle Condition in minimalist terms. Henceforth, I will assume the following version of the Strict Cycle Condition, which incorporates the main features of all the precedessors discussed in the previous subsections and at the same time evades the objections that have been raised (but see chapter 3 below for a slight modification).

(74) *Strict Cycle Condition* (second revision):
No rule can target a position that is dominated by a cyclic node and does not belong to the minimal residue of the head of this cyclic node.

The notion of minimal residue has already played a role in the definition of Last Resort in subsection 3.4 (recall (43)); it is repeated here in (75):

[36] Note that more or less the same kinds of problems arise under alternative derivations of the Strict Cycle Condition in the minimalist program; see, e.g., Kitahara (1995, 52-63), Zwart (1993, 26; 1994, 286), and, in particular, Chomsky's (1995, 234) requirement that strong features must be checked at the root. I will return to this latter condition in subsection 7.1 of chapter 3.

(75) *Minimal Residue* ('Checking Domain'):
The minimal residue of a head X is the smallest set M of categories α such that:
a. X-Adj, SpecX, and XP-Adj is in M.
b. If α is in M, then α-Adj is in M.

The notion of cyclic node in (74) still has to be defined. Here I will assume that no XP is per se a cyclic node; rather, an XP becomes a cyclic node if movement into its minimal residue (checking domain) takes place. This is the gist of the cycle:

(76) *Cycle*:
An XP becomes a cyclic node in the derivation iff movement to the minimal residue of its head takes place.

Given the notions of minimal residue and cyclic node in (75) and (76), respectively, it follows from the Strict Cycle Condition that after movement into the minimal residue of some XP α has taken place, any subsequent movement must end up either in the minimal residue of α again or in a higher position. Clearly, the Strict Cycle Condition in (74) straightforwardly rules out acyclic derivations of sentences involving *wh*-island effects at S-structure. Furthermore, (74) now makes it possible to subsume head movement under the Strict Cycle Condition. Head movement does not have to "extend" the tree built up so far anymore; the only requirement that the Strict Cycle Condition imposes on head movement is that it may not adjoin a head W to some head X if X is dominated by a cyclic node YP (Y $\neq$ X) – but if W is adjoined to X and XP is already a cyclic node, this does not violate strict cyclicity per se. More generally, given (74), head movement to X, substitution in SpecX, and adjunction to XP can co-occur because they all target the minimal residue of the same head.

Turning now to the interaction of strict cyclicity and LF movement, we must distinguish between a pre-Spell-Out (S-structure) cycle and a post-Spell-Out (LF) cycle. That is, after Spell-Out, the characterization of cyclic nodes that results from overt movement is forgotten, and everything starts anew. With respect to the illegitimate derivation (73) (repeated here as (77)) of the Korean example (71-b), this means that at the point where the first LF operation (i.e., long-distance movement of *wae*) takes place, there are no cyclic nodes present anymore. Form Chain applying to *wae* then turns the matrix CP into a cyclic node, and subsequent raising of the *wh*-argument *muŏs-ŭl* to the embedded SpecC position violates the Strict Cycle Condition (74), as desired (the embedded SpecC position is not part of the matrix C node's minimal residue).

(77) a. *S-structure*:
[$_{CP}$ – nŏ-nŭn [$_{CP}$ – Suna-ka muŏs-ŭl$_2$ wae$_1$ sa-ss-nŭnchi] alko-sip-ni] ?
b. *Form Chain applies long-distance to 'wae$_1$'*:
[$_{CP}$ wae$_1$ nŏ-nŭn [$_{CP}$ t$'_1$ Suna-ka muŏs-ŭl$_2$ t$_1$ sa-ss-nŭnchi] alko-sip-ni] ?
c. *Form Chain applies to 'muŏs-ŭl$_2$'*:
[$_{CP}$ wae$_1$ nŏ-nŭn [$_{CP}$ muŏs-ŭl$_2$ Suna-ka t$_2$ t$_1$ sa-ss-nŭnchi] alko-sip-ni] ?

4.2. Successive Cyclicity

In this section, I will present evidence that in addition to the Strict Cycle Condition, yet another constraint on rule interaction seems to be necessary, one that restricts successive-cyclic application of rules to a single item. The constraint I will argue for is the Principle of Unambiguous Binding (PUB), which rules out cases of improper movement.

4.2.1. Super-Raising in English

Consider a standard case of super-raising in English, like (78).

(78) *John$_1$ seems [$_{CP}$ that [$_{IP}$ it is likely [$_{IP}$ t$_1$ to win]]]

Here, the NP *John* has undergone Case-driven movement out of its non-Case position in the most deeply embedded IP to the matrix SpecI position. To ensure the illformedness of super-raising constructions like the one at hand, various derivations must be excluded. One derivation looks as indicated in (78); it involves movement of NP$_1$ to the matrix SpecI position in one fell swoop. This derivation violates the Barriers Condition: CP is a barrier for t$_1$ because t$_1$ is dominated by C', and the antecedent of t$_1$ is outside of CP. A second relevant derivation of (78) involves successive-cyclic A-movement of NP$_1$ via the embedded SpecI position, with subsequent insertion of the expletive pronoun *it*. Since the insertion operation targets a position that is dominated by a cyclic node – the matrix IP – without belonging to the matrix I's minimal residue, this derivation violates the Strict Cycle Condition, in addition to the Barriers Condition. But what about a third derivation of (78) in which *John* undergoes movement to the embedded SpecC position prior to undergoing raising to the matrix SpecI position, as in (79)?[37]

(79) *John$_1$ seems [$_{CP}$ t$'_1$ that [$_{IP}$ it is likely [$_{IP}$ t$_1$ to win]]]

Here, locality requirements for both the first (A-bar movement) and the second (A-movement) step are respected (see, however, Branigan (1992, 16-17) for a different view, based on a more complicated formulation of locality theory); and so is the Strict Cycle Condition. Hence, it seems that the illformedness of a derivation of the kind in (79) is not yet guaranteed by the standard constraints adopted in the minimalist program that have been discussed up to now, and additional assumptions are called for (cf. Takahashi (1994, 109-114)).

---

[37] Of course, given the concept of Form Chain, this terminology must be understood metaphorically. According to this view, what happens in (79) is not a combination of two movement operations, but rather one operation of Form Chain that consists of moving *John* to the matrix SpecI position and inserting an intermediate trace in the embedded SpecC position.

One strategy for ruling out a derivation like (79) is based on the idea that two distinct movement types interact in an illicit manner in this case, thereby creating an “improper movement” effect. The prohibition against such improper movement is derived by invoking principles of the binding theory in May (1979) and Chomsky (1981): Assuming that traces which are locally A-bound qualify as anaphors (i.e., they obey principle A of the binding theory) and that traces which are locally A-bar bound count as variables (subject to principle C), it emerges that t_1 in (79) is a variable and t_1' an anaphor. Whereas nothing is wrong with the binding of the anaphoric trace t_1' in (79), the variable t_1 violates principle C: it is bound by *John*, which occupies an A-position.

Note that this approach correctly predicts that the reverse application of A- and A-bar movement operations is legitimate. Thus, sentences like (80) do not involve improper movement under this analysis because t_1 is an anaphoric trace (being locally A-bound) that fulfills principle A, and t_1' and t_1'' are variables (being locally A-bar bound) that fulfill principle C:

(80) Who$_1$ do you think [$_{CP}$ t_1'' [$_{IP}$ t_1' was kissed t_1]] ?

Although invoking binding-theoretic principles for traces successfully accounts for improper movement effects with super-raising, closer inspection suggests that the phenomenon might be more general, i.e., that there are cases of improper movement that do not easily lend themselves to an analysis in terms of principles of binding theory. One such case, which is discussed at length in Sternefeld (1990) and Müller & Sternefeld (1990; 1993), concerns the prohibition against long-distance scrambling from finite clauses in German.

4.2.2. Long-Distance Scrambling in German

As observed by Ross (1967), scrambling in languages like German is strictly local in the sense that a finite clause cannot be crossed by this movement operation (see also subsection 2.5 of chapter 1). A relevant pair of examples that shows the clause-boundedness of scrambling in German is given in (81):

(81) a. dass Antje sagt [$_{CP}$ dass Hygrometer$_1$ niemand t_1 mag]
that Antje says that hygrometers no-one likes
b. *dass Antje Hygrometer$_1$ sagt [$_{CP}$ t_1' dass niemand t_1 mag]
that Antje hygrometers says that no-one likes

Clearly, an analysis that treats the construction in (81-b) as an instance of improper movement, on a par with the super-raising construction in (79), suggests itself. And indeed, if one assumes that scrambling in German is an instance of A-movement, the illformedness of (81-b) follows in the same way as that of the super-raising derivation in (79) (see Fanselow (1990)). However, it has often been argued in the literature that scrambling in German is to be analyzed as an instance of A-bar movement, not A-movement (cf. Stechow & Sternefeld (1988), Vikner (1990; 1994), Wilder (1989), Webelhuth (1987), Sternefeld (1990; 1991), Müller & Sternefeld (1990; 1994), Frank, Lee & Rambow (1992), Grewendorf (1995), and Müller (1995)). Assuming this to be correct, a problem arises: The initial trace t_1 in (81-b), contextually classified as a variable because it is locally A-bar bound (by the intermediate trace t_1' in SpecC), does not violate principle C because its chain antecedent *Hygrometer* ('hygrometers') occupies an A-bar, rather than an A-position. Thus, the binding-theoretic approach to improper movement in super-raising constructions does not carry over to apparently similar cases of improper movement in long-distance scrambling constructions in German.

In light of this, it is suggested in Müller & Sternefeld (1990; 1993) that a more unified approach to improper movement is called for, one that does not rely on binding theory conditions but at the same time capitalizes on the insight behind May's (1979) and Chomsky's (1981) approach to super-raising – that in these cases it is the intermediate trace in the embedded SpecC position that is illegitimate; it must be absent, thereby producing an effect akin to a *Wh*-Island Condition violation (i.e., SpecC is "inaccessible" in these contexts). The approach to improper movement developed in Müller & Sternefeld (1990; 1993) (see also Sternefeld (1991) and Müller (1995)) centers around the Principle of Unambiguous Binding, or PUB.

4.2.3. Unambiguous Binding

The Principle of Unambiguous Binding (PUB) is formulated in Müller & Sternefeld (1993, 461) as a filter that applies at S-structure and at LF:[38]

(82) *Principle of Unambiguous Binding* (PUB):
A variable that is α-bound must be β-free in the domain of the head of its chain (where α and β refer to different types of positions).

Suppose, as before, that variables – i.e., the kinds of traces that must be bound unambiguously – are defined as locally A-bar bound traces; suppose furthermore, again as before, that among the different types of positions recognized by (82) are SpecI (the landing site of A-movement), SpecC (the landing site of *wh*-movement), SpecTop (the landing site of topicalization), and left-adjunction sites of VP (IP) (the landing sites of scrambling).

---

[38]Evidently, this is not compatible with the view that S-structure does not exist as a syntactic level. As shown in Müller & Sternefeld (1996), the PUB can equally well be formulated as a derivational constraint, as in (i):

(i) *PUB* (derivational version):
There is no movement from a position of type α to a position of type β (where α is a derived A-bar position).

Although there are some subtle differences between the two formulations, I will abstract away from this issue in what follows, sticking to the representational version, mainly for reasons of simplicity.

Under these assumptions, the PUB in (82) predicts that A-bar movement to a certain kind of position (say, to SpecC) cannot be followed by another movement step that ends up in a different kind of position (say, SpecI or XP-Adj). Thus, improper movement effects as they show up in super-raising and long-distance scrambling constructions like (79) and (81-b) are derived as violations of the unambiguous binding requirement in (82).

In (79), *wh*-movement to SpecC is followed by A-movement to SpecI; such a mixing of movement types counts as improper:

(83) *John$_1$ seems [$_{CP}$ t$'_1$ that [$_{IP}$ it is likely [$_{IP}$ t$_1$ to win]]]

Note that the reverse combination of raising to SpecI and movement to SpecC as in (84) (or (80)) does not violate the PUB because t$_1$ is not a variable (it is locally A-bound) and t$'_1$ is bound unambiguously:

(84) [$_{CP}$ Who$_1$ [$_{IP}$ t$'_1$ was kissed t$_1$]] ?

However, a combination of *wh*-movement to SpecC followed by scrambling to VP, as in (81-b), does violate the PUB because it induces an ambiguous binding configuration with the variable t$_1$:

(85) *dass Antje Hygrometer$_1$ sagt [$_{CP}$ t$'_1$ dass niemand t$_1$ mag]

4.2.4. *Complications*

Thus, the PUB correctly rules out derivations of super-raising in English and long-distance scrambling in German that involve improper movement via the embedded SpecC position. However, in order to fully derive the impossibility of the respective constructions, more must be said. For instance, it is not a priori clear why certain derivations in which super-raising and long-distance scrambling constructions do not involve improper movement are completely ungrammatical, in contrast to, e.g., argument extractions from *wh*-islands in English, as in (i-a) of footnote 19 (cf. subsection 3.2.5), which is repeated here:

(86) ??Which car$_1$ did John tell you [$_{CP}$ how$_2$ (t$'_1$) to fix t$_1$ t$_2$] ?

As a case in point, consider the following, more intricate (but nonetheless ungrammatical) example involving super-raising:

(87) *John$_1$ seems [$_{CP}$ (t$''_1$) that [$_{IP}$ it is likely [$_{IP}$ t$'_1$ to have been kissed t$_1$]]]

(87) differs minimally from the kind of super-raising construction discussed so far in that the most deeply embedded verb is passivized, so that an additional trace in the direct object position of *kissed* in present. Again, if movement proceeds via SpecC, establishing an intermediate trace t$''_1$ in that position, it counts as improper due to the PUB (t$'_1$ being a variable that is bound from two different types of position, viz. SpecC and SpecI). If, on the other hand, the embedded SpecC position is skipped over, strong

ungrammaticality can be derived only if the trace in SpecI (t'_1) is not deleted on the way to LF – for this is the only "offending" trace that violates locality (the Barriers Condition). Thus, it seems that this intermediate trace of A-movement behaves like an adjunct trace and not like other argument traces (with antecedents in A-bar position) in that respect. Indeed, something along these lines is explicitly assumed by Chomsky (1986), Cinque (1990), Rizzi (1990), and Manzini (1992), among others. In contrast, in the framework adopted here, it follows automatically that t'_1 in (87) cannot be deleted on the way to LF, just like adjunct traces. Recall the notion of legitimate LF objects from Chomsky & Lasnik (1993, 546) that we already encountered in the course of the discussion of argument/adjunct asymmetries in subsection 3.2.5:

(88) *Legitimate LF objects*:
   a. Uniform chains of the type C = $<\alpha_1,...,\alpha_n>$, where all members are either in an A-bar position or in an A-position.
   b. Operator-variable constructions of the type $<\alpha,\beta>$, where α is in an A-bar position and β heads a uniform chain.

Evidently, the chain $<$*John*$_1$,t'_1,$t_1>$ in (87) contains only members in A-position; hence, it qualifies as a uniform chain in the sense of (88-a). Therefore, LF deletion of t'_1 (which would eliminate the violation of locality principles at LF and thereby produce a derivation that has the status of a Subjacency effect) is not required for convergence and is consequently blocked by the Fewest Steps condition. In sum, the illformedness of super-raising constructions follows from the interaction of the PUB, the Barriers Condition, and the Fewest Steps condition.[39]

The question arises of whether this account of illegitimate super-raising carries over to illicit long-distance scrambling from finite clauses in German. We have seen that a derivation via SpecC, as in (85), counts as improper because it causes a PUB violation. But what about a derivation like (89), where no intermediate trace is generated?

(89) *dass Antje Hygrometer$_1$ sagt [$_{CP}$ – dass niemand t_1 mag]
   that Antje hygrometers says that no-one likes

This derivation clearly violates the Barriers Condition: Any CP is a barrier for an item that is included in C'. Thus, movement from t_1 to a VP-adjoined position in the matrix clause illegitimately crosses a barrier.

Consider finally a derivation of long-distance scrambling in German where intermediate traces are generated in scrambling positions (and SpecC is still left unused), as in (90):

(90) *dass Antje Hygrometer$_1$ sagt [$_{CP}$ – dass t''_1 niemand t'_1 t_1 mag]

[39] Note that if t'_1 is not generated in the first place, t_1 will be the "offending" trace, violating locality and producing strong ungrammaticality at LF.

that Antje hygrometers says that no-one likes

This derivation does not violate the PUB because all traces are unambiguously bound by items in scrambling positions. As before, the Barriers Condition is violated: The intermediate trace t_1'', adjoined to the embedded IP, is separated from its chain antecedent by the intervening CP barrier. Therefore, this trace is marked * in the step of the derivation that generates it. This suffices to explain a Subjacency-like effect. However, in order to derive full ungrammaticality in the framework of Chomsky & Lasnik (1993), it must be guaranteed that t_1'' is not deleted on the way to LF; and, in contrast to what is the case in parallel constructions involving successive-cyclic super-raising, this does not yet follow under present assumptions – deletion of intermediate traces seems to be required here, so as to turn the scrambling chain into a legitimate LF object in the sense of (88).

This problem is solved by the principle of Full Representation that is proposed in Müller & Sternefeld (1993, 478). From this principle it follows (among other things) that traces of scrambling chains must be present at LF, i.e., cannot be deleted. For the time being, I will leave it at that, noting that scrambling chains as they arise due to Full Representation necessitate a revision of Chomsky & Lasnik's (1993) requirements for legitimate LF objects. That said, let me now turn to a third illustration of how the PUB constrains successive-cyclic movement.

4.2.5. Wh-Movement vs. Topicalization

As noted by Fanselow (1987), Bayer (1990a), Müller & Sternefeld (1993), and d'Avis (1996), among others, there is a striking asymmetry between *wh*-movement and topicalization from a *wh*-island in German. Topicalization of an argument from a *wh*-island typically induces weak (Subjacency-like) ungrammaticality (cf. (91-a)), and topicalization of an adjunct in the same context is strictly prohibited (cf. (91-b)):

(91) a. ??[$_{\text{CP}}$ – C [$_{\text{TopP}}$ Radios$_1$ weiß ich nicht [$_{\text{CP}}$ wie$_2$ [$_{\text{TopP}}$ t_1' man t_2 t_1
radios$_{acc}$ know I not how one
repariert]]]]
fixes

b. *[$_{\text{CP}}$ – C [$_{\text{TopP}}$ Deswegen$_2$ weiß ich nicht [$_{\text{CP}}$ was$_1$ [$_{\text{TopP}}$ t_2' der Fritz t_2
therefore know I not what$_{acc}$ ART Fritz
t_1 repariert hat]]]]
fixed has

Thus, this movement type behaves more or less as one would expect it to under standard approaches to *wh*-islands. In contrast, *wh*-movement from a *wh*-island is strictly prohibited in German not only with adjuncts (cf. (92-b)), but also with arguments; as shown in (92-a), argument extraction is much worse than the corresponding topicalization case in (91-a).

(92) a. ?*[$_{\text{CP}}$ Was$_1$ C [$_{\text{TopP}}$ – weißt du nicht [$_{\text{CP}}$ wie$_2$ [$_{\text{TopP}}$ – man t$_2$ t$_1$
what$_{acc}$ know you not how one
repariert]]]] ?
fixes

b. *[$_{\text{CP}}$ Wie$_2$ C [$_{\text{TopP}}$ – weißt du nicht [$_{\text{CP}}$ was$_1$ [$_{\text{TopP}}$ – man t$_2$ t$_1$
how$_2$ know you not what$_{acc}$ one
repariert]]]] ?
fixes

Note first that if the examples in (91) and (92) are derived in such a way that embedded *wh*-movement follows long-distance movement, the Strict Cycle Condition is violated as shown above, yielding strong ungrammaticality, by assumption. Thus, suppose that these sentences are derived in a cyclic fashion, with embedded *wh*-movement preceding long-distance extraction. In this case, the embedded CP is a barrier due to the fact that the embedded SpecC position is filled by the *wh*-phrase, and extraction violates the Barriers Condition (1) in the derivation. This suffices to predict strong ungrammaticality in the cases of adjunct extraction in (91-b) and (92-b). The intermediate status of the example involving argument topicalization in (91-a) is also accounted for – due to a Barriers Condition violation, the trace t'_1 is marked * in the derivation, thereby inducing weak ungrammaticality, but t'_1 is deleted on the way to LF, so that strong ungrammaticality is avoided. However, the clear contrast between topicalization of an argument from a *wh*-island as in (91-a) and *wh*-movement of an argument from a *wh*-island as in (92-a) is not yet predicted.

The idea pursued in Müller & Sternefeld (1993) in light of this evidence is the following: Obviously, SpecC in (91) and (92) is not accessible for long-distance movement because it is filled and thus creates a violation of the Barriers Condition. However, similar conclusions can be drawn in the case of SpecTop – if this position is skipped, too, then yet another violation of the Barriers Condition will occur. It is reasonable to postulate that such an additional violation can be held responsible for the more degraded status of examples involving argument extraction across a *wh*-island. Thus, the task is to ensure that SpecTop can qualify as an escape hatch for long-distance movement in (91-a), but not in (92-a), resulting in a weak violation in the former and a stronger violation in the latter.

This result is achieved by the PUB. As shown in (93-a) vs. (93-b), *wh*-movement via SpecTop is classified as improper by the PUB, whereas successive-cyclic topicalization via SpecTop is not.

(93) a. ?*[$_{\text{CP}}$ Was$_1$ C [$_{\text{TopP}}$ – ... [$_{\text{CP}}$ wie$_2$ [$_{\text{TopP}}$ t'_1 man t$_2$ t$_1$ repariert]]]] ?
b. ??[$_{\text{CP}}$ – C [$_{\text{TopP}}$ Radios$_1$ ... [$_{\text{CP}}$ wie$_2$ [$_{\text{TopP}}$ t'_1 man t$_2$ t$_1$ repariert]]]]

Of course, this analysis also raises a number of problems. For one thing, the partition of the clausal periphery into CP and TopP, together with the PUB, requires special assumptions for ordinary cases of successive-cyclic *wh*-movement and topicalization, as in (94):

(94) a. [$_{CP}$ Wie$_1$ C [$_{TopP}$ – Top meinst du [$_{CP}$ t$'_1$ dass [$_{TopP}$ – Top das t$_1$
how think you that that$_{nom}$
passiert ist]]]] ?
happened is

b. [$_{CP}$ – C [$_{TopP}$ Dieses Buch$_1$ denke ich [$_{CP}$ – dass [$_{TopP}$ t$'_1$ ich t$_1$ kaufen
this book$_{acc}$ think I that I buy
sollte]]]]
should

The only derivations of (94-a) and (94-b) that are compatible with the PUB are those that have intermediate traces in the positions indicated here, i.e., in the embedded SpecC position in the *wh*-movement example in (94-a) and in the embedded SpecTop position in the topicalization example in (94-b). If nothing more is said, we would therefore expect violations of the Barriers Condition in these sentences, resulting in strong illformedness with adjunct movement in (94-a) and weak illformedness with argument movement in (94-b), contrary to fact (in (94-a), both the embedded TopP and the matrix TopP should qualify as barriers, and in (94-b), the embedded CP should be a barrier).

Faced with this problem, Müller & Sternefeld (1990; 1993) suggest that CP and TopP count as "matching projections" (cf. Haider (1988) for a related concept) in (94) in the sense that they behave like a single XP for the purposes of locality constraints in these examples. The matching of two projections is, according to this view, possible only if one specifier is radically empty (i.e., does not contain lexical material or traces) – this ensures that matching is not an option in the examples in (91) and (92), where both CP and TopP must be visible for the Barriers Condition, as seen above.

The PUB has many more consequences for successive-cyclic extraction from clauses under the CP/TopP hypothesis.[40] However, I will not dwell on this issue any further here (see Müller & Sternefeld (1990; 1993; 1996) for more discussion) and turn now to the question of how the cyclicity constraint PUB and the economy constraints Fewest Steps and Last Resort interact with each other.

[40] To name just one, *wh*-movement of an argument across a *wh*-island in English should have the same status as German examples of the type (92-a); i.e., it should involve two violations of the Barriers Condition, yielding a severely degraded result (that is, however, still somewhat better than the corresponding case of adjunct extraction). Indeed, it is argued in Müller & Sternefeld (1990; 1993) that this result might be correct for finite *wh*-islands in English, which are typically strict, in contrast to non-finite *wh*-islands (see Chomsky (1986) and Frampton (1990)). As far as non-finite *wh*-islands as in (86) are concerned, it is proposed that here the TopP projection is "impoverished" to an extent that permits *wh*-items to use SpecTop without incurring a PUB violation. On this view, the weak islandhood of infinitival *wh*-clauses for *wh*-movement in English can be treated on a par with the weak islandhood of finite *wh*-clauses for topicalization in German; cf. (91-a).

4.2.6. The Interaction of the PUB and Economy Constraints

In this subsection, I would like to address the question of whether the effects of improper movement discussed in the preceding subsections really require a specific constraint like the PUB or can be made to follow from the economy constraints Fewest Steps and Last Resort in their standard forms. I have highlighted three cases involving improper movement: (a) super-raising in English, (b) long-distance scrambling in German, and (c) a *wh*-movement/topicalization asymmetry in the case of *wh*-islands in German. The three relevant derivations are repeated here in that order (see (83), (85), and (93-a)):

(95) a. *John$_1$ seems [$_{CP}$ t$'_1$ that [$_{IP}$ it is likely [$_{IP}$ t$_1$ to win]]]
b. *dass Antje Hygrometer$_1$ sagt [$_{CP}$ t$'_1$ dass niemand t$_1$ mag]
that Antje hygrometers says that no-one likes
c. ?*[$_{CP}$ Was$_1$ C [$_{TopP}$ – ... [$_{CP}$ wie$_2$ [$_{TopP}$ t$'_1$ man t$_2$ t$_1$ repariert]]]] ?

As can easily be verified, the Fewest Steps condition does not rule out the derivations in (95). In all three cases, only one operation of Form Chain applies to the moved item; there is nothing inherent in that operation that precludes the insertion of the "offending" intermediate traces that give rise to improper movement configurations. Thus, since Fewest Steps per se does not say anything about possible positions for the insertion of intermediate traces, it cannot block the derivations in (95) in favour of competing derivations which do not contain improper movement.

Note next that the PUB effects in (95) do not follow from Last Resort either. In all cases, long-distance movement (to SpecI, VP-Adj, and SpecC, respectively) is triggered by the need to check a morphological feature ([scr], [Case], and [+wh]). Like Fewest Steps, Last Resort does not per se impose a restriction on the insertion of intermediate traces by Form Chain, and hence does not bar improper movement configurations.

More generally, then, the conclusion can be drawn that a constraint on successive cyclicity like the PUB appears to be necessary, at least under the standard formulations of Last Resort and Fewest Steps. Unfortunately, a redundancy nevertheless arises, but in the opposite direction. It turns out that the PUB can also do some of the work of economy constraints – as noted in Müller & Sternefeld (1996), the PUB directly excludes both *wh*-topicalization and *wh*-scrambling, as discussed in subsections 3.2.1 and 3.2.2 above. Consider first an example that illustrates the prohibition against *wh*-topicalization in English. (17-c) is repeated here:

(96) *Who$_1$ t$_1$ said [$_{CP}$ that [$_{TopP}$ who$_2$ [$_{IP}$ John likes t$_2$]]] ?

At LF, *who*$_2$ is raised from the embedded SpecTop position to the matrix SpecC position, where it is then adjoined. Thus, a derivation that generates the LF representation in (97) via the S-structure representation in (96) violates Fewest Steps because there is a competing derivation that is more economical (in this derivation, *who*$_2$ stays in situ at S-structure). However, a derivation yielding (97) via (96) also involves ambiguous binding of the inital trace t_2. Hence, such successive-cyclic movement violates the PUB and counts as improper:

(97) *[$_{\text{SpecC}}$ who$_1$ who$_2$] t_1 said [$_{\text{CP}}$ that [$_{\text{TopP}}$ t'_2 [$_{\text{IP}}$ John likes t_2]]]

Basically the same goes for the prohibition against *wh*-scrambling in German. The relevant example is (19-c), repeated here as (98):

(98) *Wie$_1$ hat [$_{\text{IP}}$ was$_2$ [$_{\text{IP}}$ der Fritz t_1 t_2 repariert]] ?
how has what$_{acc}$ ART Fritz fixed

As before, the *wh*-phrase that has undergone optional movement at S-structure (i.e., *was*$_2$ ('what')) raises to the matrix SpecC position at LF; and again, the Fewest Steps condition rules out such a derivation in favour of a more economical derivation in which *was*$_2$ is in situ at S-structure. However, overt *wh*-scrambling also induces an improper movement configuration (reducible to the PUB) – at LF, t_2 is bound ambiguously from a left-adjoined IP (i.e., scrambling) position and a right-adjoined SpecC (i.e., *wh*-) position:

(99) *[$_{\text{SpecC}}$ [$_{\text{SpecC}}$ wie$_1$] was$_2$] ... [$_{\text{IP}}$ t'_2 [$_{\text{IP}}$ ... t_2]]

I take it that such a redundancy is not desirable; before I turn to its resolution in section 6, I will highlight further problems with the present analysis in the following section.

5. Problems

5.1. Conceptual Problems

The approach developed so far suffers from a number of drawbacks. A first conceptual problem has just been mentioned: There is a redundancy between cyclicity constraints (the PUB) and economy constraints (Fewest Steps) as far as the ban on *wh*-topicalization in English and *wh*-scrambling in German is concerned. A second problem is conceptual in nature, too. It concerns the notion of Form Chain. As shown in subsection 3.2.4 above, the concept of Form Chain as in (100) (= (32)) as a replacement for Move α appears to be necessary in order to reconcile locality requirements with the transderivational economy condition Fewest Steps.

(100) *Form Chain*:
Move α to its target position and insert intermediate traces in positions required by locality conditions.

Unless Form Chain is adopted, the Fewest Steps condition (cf. (16), repeated here as (101)) wrongly precludes successive-cyclic movement, which is required in (102) (= (30)):

(101) *Fewest Steps*:
If two derivations D_1 and D_2 are in the same reference set and D_1 involves

fewer operations than D_2, then D_1 is to be preferred over D_2.

(102) Why$_1$ do you think [$_{CP}$ t'_1 that Mary came t_1] ?

One might assume that with the introduction of a derivational economy condition like Last Resort in Chomsky (1993; 1995), the concept of Form Chain could be dispensed with. Recall that given the formulation of Last Resort in (42) (repeated here as (103)), movement is possible only if it results in feature checking.

(103) *Last Resort*:
α is raised to a position β only if β is in the minimal residue of γ, and α enters into a checking relation with γ.

Thus, a truly successive-cyclic derivation of a sentence like (102) that involves two movement operations applying to the *wh*-phrase would have to involve two feature checking operations in which the *wh*-phrase participates, one in each SpecC position that is targeted. The numeration of a derivation of (102) in one fell swoop would, in contrast, necessarily lack the relevant features that are checked in the embedded SpecC positions; otherwise, it would not converge. Therefore, at first sight it seems as though a derivation of (102) in one fell swoop (without an intermediate trace) and a successive-cyclic derivation of (102) do not compete, due to different feature specifications in the numeration. However, this conclusion is not correct. Recall from subsection 3.2.3 that the stipulation in (104) (= (28)) had to be adopted in order to ensure that a derivation with an optional feature (like [top] or [scr]) can compete with (and ultimately be excluded by) a derivation that differs minimally in that such an optional feature is absent.

(104) *Identity of Numerations*:
Two numerations are identical if they differ only in the feature specification on some head.

Thus, given (104), a derivation of (102) in one fell swoop would still compete with a successive-cyclic derivation of (102) in which all *wh*-movement steps are triggered by feature checking operations, in accordance with Last Resort. Hence, without recourse to Form Chain, the Fewest Steps condition would still incorrectly predict derivations in one fell swoop to block successive-cyclic derivations. In conclusion, Last Resort does not render Form Chain superfluous under present assumptions, as far as the Fewest Steps condition is concerned.

What is more, it seems that Last Resort itself requires the concept of Form Chain. As noted above, a derivation of (102) involving two movement operations applying to the *wh*-phrase would require two instances of feature checking with the *wh*-phrase, one in the embedded [–wh] SpecC position and one in the matrix [+wh] SpecC position. More generally, we are led to the assumption that a *wh*-phrase that crosses *n* [–wh] SpecC positions on its way to its ultimate [+wh] target position is involved in *n+1* feature checking operations in SpecC positions; clearly, the features that are checked in these positions must be distinct from the *wh*-feature (the positions are marked [–wh]). Indeed, something along these lines has occasionally been suggested in recent work (cf. Collins (1994a; 1997, ch. 5), Groat & Ferguson (1994), Fanselow & Mahajan (1996), and Sabel (1996); see also the discussion in Chomsky (1995, 250 & note 29; 267; and 291)), and there may be direct evidence for such an assumption in languages that show overt reflexes of successive-cyclicity, such as Ewe or Modern Irish. However, I assume that the postulation of these kinds of feature checking operations in intermediate SpecC positions is poorly motivated in languages like English or German and should be avoided if possible. But if it is avoided, this implies that Last Resort requires the concept of Form Chain – according to this view, the *wh*-phrase in sentences like (102) is moved to the matrix [+wh] SpecC position (where it participates in feature checking) in one fell swoop, and intermediate traces are inserted by the operation Form Chain in the embedded SpecC positions (without further checking operations taking place).

Summarizing, it seems that economy constraints strongly suggest a concept like Form Chain. Ideally, however, one would like to be able to dispense with such an artifact of the theory. Moreover, it is well known that Form Chain raises a theory-internal problem in the minimalist program – it violates strict cyclicity because the insertion of intermediate traces after the movement operation involves referring back to a proper subpart of the tree that has been created.[41] Therefore, I conclude that Form Chain should be abandoned.

Summarizing, there are two conceptual problems with the approach developed so far: On the one hand, it creates a redundancy (the PUB and Fewest Steps both rule out *wh*-topicalization and *wh*-scrambling), and on the other hand, it necessitates the problematic concept of Form Chain as a replacement for Move α. In the following subsection, I turn to an empirical problem raised by the analysis.

5.2. An Empirical Problem

5.2.1. Evidence for Modifying the Notion of Reference Set

Recall first that the Fewest Steps condition rules out *wh*-topicalization in English, as in (105):

(105) *Who$_1$ t$_1$ said [$_{CP}$ that who$_2$ [$_{IP}$ John likes t$_2$]] ?

As observed in Müller & Sternefeld (1996, 482), the same prohibition applies in German. Here, too, *wh*-topicalization (as in (106-c)) is not permitted, although there is no general prohibition against embedded topicalization in this context (cf. (106-b)) and although multiple questions can in principle be formed in which a *wh*-phrase in an

[41] Accordingly, Chomsky (1993, 46) notes in this connection that “a modification [of strict cyclicity] is necessary for the case of successive-cyclic movement, interpreted in terms of the operation Form Chain.”

embedded clause takes matrix scope (cf. (106-a)).

(106) a. Wer$_1$ sagte t$_1$ [$_{CP}$ dass [$_{IP}$ Fritz wen$_2$ gesehen habe]] ?
who said that Fritz whom$_{acc}$ seen has$_{subj}$

b. Wer$_1$ sagte t$_1$ [$_{CP}$ den Fritz$_3$ habe [$_{IP}$ er t$_3$ gesehen]] ?
who said ART Fritz$_{acc}$ has$_{subj}$ he seen

c. *Wer$_1$ sagte t$_1$ [$_{CP}$ wen$_2$ habe [$_{IP}$ Fritz t$_2$ gesehen]] ?
who said whom$_{acc}$ has$_{subj}$ Fritz seen

Obviously, whatever explanation is given for the illformedness of the English example in (105) should carry over to the German example in (106-c). However, it emerges that the Fewest Steps condition does not rule out (106-c) under present assumptions.[42] To see this, consider the (simplified) LF representation which is generated by derivation D$_1$ that yields the S-structure representation in (106-c):[43]

(107) *[SpecC Wer$_1$ wen$_2$] sagte t$_1$ [$_{CP}$ t$'_2$ [$_{IP}$ Fritz t$_2$ gesehen habe]] ?

Assuming complementizer deletion at LF (cf. Lasnik & Saito (1984; 1992), Chomsky (1986), among many others), a derivation D$_2$ that generates the S-structure representation in (106-a) will end up with the same LF representation. Furthermore, assuming deletion of expletive *es* as it may appear in the SpecV/2 position in German V/2 clauses, a derivation D$_3$ that yields the S-structure representation in (108) will also have the same LF representation:[44]

(108) ?Wer$_1$ sagte t$_1$ [$_{CP}$ es habe [$_{IP}$ Fritz wen$_2$ gesehen]] ?
who said EXPL has$_{subj}$ Fritz whom$_{acc}$ seen

Thus, one might hope that because they yield identical LF representations, the derivations D$_2$ and D$_3$ generating the S-structure strings in (106-a) and (108) compete with the derivation D$_1$ that generates the ill-formed *wh*-topicalization structure in (106-c), and that D$_1$ is blocked by D$_2$ and D$_3$ since it involves an additional operation applying to the *wh*-phrase *wen*$_2$ that is undone at LF.[45] Unfortunately, under the notion of reference set assumed so far (cf. (29), repeated here as (109)), this is not the case.

---

[42] As shown in Müller & Sternefeld (1996), the derivation generating (106-c) violates the PUB, just as the derivation generating (105) does.

[43] I assume here that overt V/2 movement is reconstructed at LF for semantic reasons; cf. Stechow (1992). Also see Koster (1987) for syntactic evidence that points to this conclusion.

[44] For reasons which are not clear to me, V/2 clauses in general do not appear to tolerate *wh*-in situ taking matrix scope easily in German; cf. Büring & Hartmann (1994, 60) for discussion.

[45] This would still presuppose that D$_1$ – D$_3$ are equally costly as far as other operations (V/2 movement, V/2 reconstruction, complementizer insertion, complementizer deletion, expletive insertion, expletive deletion) are concerned. For the purposes of the discussion, I will assume that this is indeed the case.

(109) *Reference Set*:
Two derivations D_1 and D_2 are in the same reference set iff they start with the same numeration, converge at LF and PF, and yield the same LF representation.

The numeration of D_1 differs from that of D_2 in that D_2 has an additional complementizer and from that of D_3 in that D_3 has an additional expletive *es*. For this reason, I would like to suggest that the notion of numeration be dispensed with in the definition of reference set and that (109) be altered as in (110):

(110) *Reference Set* (final revision):
Two derivations D_1 and D_2 are in the same reference set iff they yield the same LF output and converge at LF and PF.

Interestingly, this move does not seem to produce any negative consequences for the cases involving transderivational economy constraints that were discussed so far. Furthermore, the revised notion of reference set now ensures that *wh*-topicalization in German as in (106-c) is blocked by Fewest Steps; and it also yields the welcome result that the assumption about identity of numerations in (104) can now be dispensed with. However, a problem arises in the shape of partial *wh*-movement constructions.

5.2.2. Partial Wh-Movement

Many languages exhibit the phenomenon of partial *wh*-movement. In partial *wh*-movement constructions, the *wh*-phrase undergoes S-structure movement to a [–wh] SpecC position; to reach its ultimate LF [+wh] target position (which is occupied by a scope marker at S-structure in some languages), a second (covert) movement operation appears necessary.[46] Partial *wh*-movement systematically and freely alternates with successive-cyclic long-distance *wh*-movement in German; this is shown in (111-a) vs. (111-b) (cf. van Riemsdijk (1982), Stechow & Sternefeld (1988), McDaniel (1989), and the contributions in Lutz & Müller (1996)).

(111) a. Wen$_1$ meinst du [$_{\text{CP}}$ t$'_1$ dass sie wirklich t$_1$ liebt] ?
who$_{acc}$ think you that she really loves
b. Was$_1$ meinst du [$_{\text{CP}}$ wen$_1$ (dass) sie wirklich t$_1$ liebt] ?
[+wh] think you who$_{acc}$ that she really loves

Thus, given that the derivation D_1 generating the successive-cyclic movement structure (111-a) and the derivation D_2 generating the partial *wh*-movement structure

[46] Dayal (1994) argues that a partially moved *wh*-phrase in German and Hindi can be interpreted in its S-structure position and need not undergo a second movement operation at LF. Cf., however, Höhle (1990; 1996), Müller & Sternefeld (1996), and Beck & Berman (1996) for arguments that such an approach is not correct for German.

(111-b) both end up in the same LF representation (roughly of the form (111-a)), D_1 and D_2 are in the same reference set, according to (110). However, D_1 involves only one instance of Form Chain applying to the *wh*-phrase *wen*$_1$ ('who') (viz., overt long-distance movement plus insertion of an intermediate trace), whereas D_2 invariably involves two instances of Form Chain applying to *wen* – one before and one after the Spell-Out operation. Hence, the Fewest Steps condition incorrectly predicts partial *wh*-movement in German to be blocked by successive-cyclic long-distance movement.

Obviously, this problem would not arise under the notion of reference set in (109) because this notion depends on identity of numerations, and a derivation that includes a scope marker *was* automatically has a different numeration than a derivation in which a scope marker is not present. A cross-linguistic survey suggests, though, that partial *wh*-movement poses a problem under the notion of reference set in (109), too; so, the problem is more general.

Thus, whereas partial *wh*-movement in German requires a scope marker *was* in the [+wh] target position, the analogous construction in a language like Iraqi Arabic (cf. Wahba (1992)) often does not. Moreover, partial *wh*-movement in this language (cf. (112-b) and (112-c)) freely alternates with both successive-cyclic long-distance movement (cf. (112-a)) and *wh*-in situ at S-structure (cf. (112-d)).

(112) a. [$_{CP}$ meno$_1$ Mona raadat [$_{CP}$ t$'_1$ tijbir Su'ad [$_{CP}$ t$'_1$ tisa'ad t$_1$]]] ?
whom$_{dat}$ Mona wanted to force Su'ad to help

b. [$_{CP}$ – Mona raadat [$_{CP}$ meno$_1$ tijbir Su'ad [$_{CP}$ t$'_1$ tisa'ad t$_1$]]] ?
Mona wanted to whom$_{dat}$ force Su'ad to help

c. [$_{CP}$ – Mona raadat [$_{CP}$ – tijbir Su'ad [$_{CP}$ meno$_1$ tisa'ad t$_1$]]] ?
Mona wanted to force Su'ad whom$_{dat}$ to help

d. [$_{CP}$ – Mona raadat [$_{CP}$ – tijbir Su'ad [$_{CP}$ – tisa'ad meno$_1$]]] ?
Mona wanted to force Su'ad to help whom$_{dat}$

The derivations D_1 and D_2 which generate the S-structure representations in (112-a) and (112-d) can do with only one application of Form Chain to the *wh*-phrase *meno*$_1$ ('whom'), at S-structure or at LF, respectively. In contrast, the derivations D_3 and D_4 generating the partial *wh*-movement structures in (112-b) and (112-c) need two applications of Form Chain to *meno*$_1$. Since D_1 – D_4 share the same lexical material in the numeration and since they terminate in the same LF representation, they are expected to compete with respect to economy constraints under both definitions of reference set (i.e., (109) and (110)). For this reason, D_3 and D_4 are wrongly predicted to be blocked by D_1 and D_2.[47]

---

[47] Note that partial *wh*-movement constructions are a fairly widespread phenomenon; they show up in many more languages (sometimes with and sometimes without a scope marker), e.g., Ancash Quechua (cf. Cole (1982)), Bahasa Indonesia (cf. Saddy (1991)), Hungarian (cf. Marácz (1990)), and Palauan (cf. Georgopoulos (1991)), so a systematic approach is called for.

To solve the three problems discussed in subsections 4.2.6, 5.1, and 5.2, I will now develop a revised approach to economy constraints.

6. A Revised Approach to Economy

In this section, I would like to suggest modifications of the economy constraints Fewest Steps and Last Resort. The basic idea is that economy constraints should be made sensitive to a distinction between *movement operations* on the one hand and actual *checking operations* on the other.

6.1. Assumptions

Suppose first that the Fewest Steps condition does not count all movement operations per se; rather, it only counts feature checking operations. The revised Fewest Steps condition is given in (113):[48]

(113) *Fewest Steps* (revised):
If two derivations D_1 and D_2 are in the same reference set and D_1 involves fewer *checking* operations than D_2, then D_1 is to be preferred over D_2.

Obviously, this revision becomes interesting only if there is movement without feature checking. Therefore, suppose next that Last Resort does in fact permit movement without feature checking, but only if that movement is driven by the attempt to check a feature. More precisely, I will assume that an item α can be "fooled" into moving to a certain position if that is a typical checking position for an unchecked feature of α.[49] The revised Last Resort condition is shown in (114):

(114) *Last Resort* (revised):
α is raised to a position β only if β is a *typical checking position* for an unchecked morphological feature of α.

---

[48] As it stands, the revised Fewest Steps condition no longer says anything about syntactic operations that do not involve movement, like deletion, insertion, etc. Thus, in order to maintain the approach to argument/adjunct asymmetries developed in Chomsky & Lasnik (1993) (cf. subsection 3.2.5 above), which crucially relies on the idea that uneconomical deletion is prohibited by Fewest Steps, the condition must be reformulated. One possibility would be to assume that Fewest Steps does not count only *checking* operations (as in (113)), but rather, say, ζ-operations, where ζ stands for checking, deletion, insertion, etc., but not for movement per se. Nevertheless, for reasons of simplicity and perspicuity, I will adopt the version of Fewest Steps in (113) in what follows, keeping in mind that a more careful definition would have to proceed along the lines just sketched.

[49] The idea of fooling an item into movement is due to Collins (1994a, 17). According to Collins' formulation of Last Resort (called "Greedier"), however, a *wh*-phrase that is fooled into moving to an intermediate SpecC position in the course of successive-cyclic movement still has to participate in feature checking in that position, even if feature checking does not yet affect the [+wh] feature. As will become clear shortly, this is not the case in the approach suggested here.

As far as the notion of "typical checking position" is concerned, it simply refers to the regular checking position for a given feature, with the qualification that no actual feature checking has to take place. Thus, I will assume that the typical checking position for the feature [+wh] is a position in the minimal residue of C; that the typical checking position for the feature [Case] is a substitution position in the minimal residue of a head in the extended projection of the verb; that the typical checking position for the feature [scr] is an adjoined position in the same domain; and that the typical checking position for the feature [top] is a substitution position in the minimal residue of Top.

Based on these assumptions, I will now show that the three problems raised by the earlier approach (relating to a redundancy between the PUB and Fewest Steps, the concept of Form Chain, and the existence of partial *wh*-movement) can all be overcome. I will begin with the concept of Form Chain, turn to partial *wh*-movement after that, and finally address the derivation of PUB effects.

6.2. Form Chain and Economy

Consider again a standard example involving successive-cyclic long-distance *wh*-movement of an adjunct, as in (115) (= (102)):

(115) Why$_1$ do you think [$_{CP}$ t$'_1$ that Mary came t$_1$] ?

It turns out that the present system of economy constraints systematically permits successive-cyclic movement without recourse to a concept like Form Chain. Both of the movement operations in (115) that apply to *why*$_1$ end up in a typical checking position for the unchecked [+wh] feature of *why*$_1$ (viz., in SpecC), even though only the second instance of *wh*-movement involves actual feature checking. Hence, Last Resort as defined in (114) is respected. Similarly, Fewest Steps is respected (cf. (113)). The reason is that although (115) involves two instances of movement applying to *why*$_1$, it involves only one instance of [+wh] feature checking. Therefore, it is no more costly than a competing derivation in which *why*$_1$ has undergone *wh*-movement in one fell swoop, thereby violating the Barriers Condition. Given this state of affairs, Fewest Steps permits us to choose either the successive-cyclic derivation in (115) or the competing derivation in one fell swoop, and since the latter violates the Barriers Condition while the former does not, the derivation with successive-cyclic movement is chosen, and the sentence is correctly predicted to be well formed.[50]

[50] One might assume that the revised Last Resort condition in (114) poses a problem with respect to Procrastinate, since the the first movement step of the *wh*-phrase in (115) seems to be motivated by the need to check a *weak* [+wh] feature – recall that it is the [+wh] feature of C that is strong in English, and not the [+wh] feature of a *wh*-phrase. However, closer inspection reveals that this dilemma is only apparent. Last Resort as defined in (114) does not actually require movement of the *wh*-phrase; the only thing that it does is to permit such movement. Hence, it is compatible not only with the derivation in (115), but also with a derivation in which the *wh*-phrase moves only to the embedded SpecC position, or not at all. The

6.3. Wh-Movement and Economy

Consider next the case of partial *wh*-movement. The German example in (111-b) is repeated here as (116-a), together with its abbreviated LF representation in (116-b) (the following remarks carry over to partial *wh*-movement in Iraqi Arabic (cf. (112-b) and (112-c)) and in other languages):[51]

(116) a. Was$_1$ meinst du [$_{CP}$ wen$_1$ (dass) sie wirklich t$_1$ liebt] ?
[+wh] think you who$_{acc}$ that she really loves
b. [$_{SpecC}$ wen$_1$] ... [$_{CP}$ [$_{SpecC}$ (t$'_1$)] ... t$_1$]

Under the formulations of Fewest Steps and Last Resort in (113) and (114), the wellformedness of (116-a) follows directly, basically in the same way as the wellformedness of (115). The *wh*-phrase *wen*$_1$ ('who$_{acc}$') bears one [+wh] feature in the numeration. Overt *wh*-movement of *wen*$_1$ takes place to a typical checking position for the – as yet unchecked – feature [+wh], viz., to the embedded SpecC position. There is no actual feature checking in this position (the embedded C node does not bear a [+wh] feature), but this is permitted by Last Resort. The [+wh] feature on the *wh*-phrase is eventually checked off by subsequent LF raising to the matrix SpecC$_{[+wh]}$ position; evidently, this second movement operation respects Last Resort, too. Since this derivation involves two movement operations applying to *wen*$_1$, but only one instance of [+wh] feature checking, it is just as costly as a competing derivation involving successive-cyclic long-distance *wh*-movement at S-structure (compare (111-a)). Hence, Fewest Steps does not rule out either of these two options, and the two constructions alternate freely.[52]

---

choice among these derivations is then made by the requirement of PF convergence – only if *wh*-movement applies twice, checking the strong [+wh] feature of the matrix C node and yielding (115), will the derivation converge.

[51] I assume here that the scope marker *was*, being an illegitimate (i.e., uninterpretable) LF object, is overwritten by the *wh*-phrase at LF via expletive replacement. Alternatively, it could be assumed that the *wh*-phrase does not replace the scope marker at LF, but adjoins to it. Given that the LF output of partial *wh*-movement constructions and successive-cyclic *wh*-movement constructions would still count as identical under such an assumption (in the technical sense relevant for the definition of reference sets), this would not yield different consequences.

[52] Although this analysis reconciles the very existence of partial *wh*-movement with the economy constraints Last Resort and Fewest Steps, it is clear that many relevant questions are left open. First, there is a problem related to Procrastinate. Assuming that scope marker insertion is a means to check the strong [+wh] feature of the matrix C node in (116-a), it is ensured that the derivation converges even though the *wh*-phrase has not yet reached its LF target position at S-structure; however, given Procrastinate, it is unclear why partial *wh*-movement is possible in the first place, and why it is indeed necessary in German, i.e., why the *wh*-phrase in (116-a) does not simply stay in situ at S-structure (as in, e.g., Iraqi Arabic). A first step towards a solution might consist in making Procrastinate sensitive only to checking operations, and not movement operations in general (as it has been done in the case of Fewest Steps), but I will not go into that here. A second open question concerns language-specific variation: As it stands, the analysis proposed here

Note that whereas the present system of economy constraints permits *wh*-movement to a [–wh] SpecC position followed by LF *wh*-movement to a [+wh] SpecC position, the case is different with overt *wh*-movement to a [+wh] SpecC position followed by LF raising to another [+wh] SpecC position. The relevant evidence comes from multiple questions such as (117), a type of example studied in Baker (1970) and much subsequent work (also cf. subsection 3.4.1 above).

(117) Who$_1$ t$_1$ wonders [$_{CP}$ where$_2$ we bought what$_3$ t$_2$] ?

As is well known, (117) is ambiguous in two ways. It can be interpreted either as a multiple embedded question with *what*$_3$ taking embedded scope or as a multiple matrix question with *what*$_3$ taking matrix scope. These two readings correspond to the LF representations of (117) in (118-a) and (118-b), respectively.

(118) a. [$_{SpecC}$ who$_1$] t$_1$ wonders [$_{CP}$ [$_{SpecC}$ [$_{SpecC}$ where$_2$] what$_3$] we bought t$_3$ t$_2$]
b. [$_{SpecC}$ [$_{SpecC}$ who$_1$] what$_3$] t$_1$ wonders [$_{CP}$ [$_{SpecC}$ where$_2$] we bought t$_3$ t$_2$]

However, a third reading is absent. In the LF representation corresponding to this reading, the *wh*-phrase *where*$_2$, which is in the embedded SpecC position at S-structure, is moved to the matrix SpecC position at LF, thereby creating a multiple matrix question, and *what*$_3$ takes embedded scope. This LF representation is shown in (119):

(119) *[$_{SpecC}$ [$_{SpecC}$ who$_1$] where$_2$] t$_1$ wonders [$_{CP}$ [$_{SpecC}$ (t$'_2$) what$_3$] we bought t$_3$ t$_2$]

Various proposals have been made in order to rule out such an LF based on the S-structure in (117). The empirical generalization commonly arrived at (cf., e.g., Chomsky (1973; 1986), Aoun, Hornstein & Sportiche (1981), and Lasnik & Saito (1984; 1992)) is that overt *wh*-movement to a [+wh] position fixes the scope of the moved *wh*-phrase. As observed by Epstein (1992), this generalization suggests an account in terms of economy, and it remains to be shown that the present system achieves that.

According to the standard view, *where*$_2$ bears one [+wh] feature in the numeration; this feature is checked off by overt *wh*-movement to the [+wh] SpecC position of the embedded clause. Hence, subsequent LF raising to the matrix [+wh] SpecC position violates Last Resort (Fewest Steps would be respected by this derivation). But suppose now that *where*$_2$ actually comes equipped with *two* [+wh] features in the numeration, the second one being optional. One might argue that such a thing is precluded anyway. However, it is at least worth noting that under the present system it is unneccessary to

---

does not explain why partial *wh*-movement is impossible in a language like English. See McDaniel (1989), Stechow & Sternefeld (1988), Fanselow & Mahajan (1996), and the contributions in Lutz & Müller (1996), where these and other problems are addressed.

stipulate a prohibition against such an unmotivated proliferation of identical features explicitly. The reason is that a derivation of (119) via (117) in which *where*$_2$ has two [+wh] features in the numeration would automatically be excluded by Fewest Steps (although Last Resort would be respected in this case); there is a competing converging derivation that yields the same LF output but does so with one fewer instance of [+wh] feature checking. In this derivation, *where*$_2$ initially has only one [+wh] feature and remains in situ at S-structure (*what*$_3$ undergoes overt *wh*-movement to the embedded SpecC position). Summarizing, the interaction of Fewest Steps and Last Resort rules out *wh*-movement at LF from a [+wh] position. In addition, a more general result is obtained: The optional addition of features that trigger movement to lexical items is severely restricted; i.e., features on items are minimized. More specifically, the following theorem can now be derived:[53]

(120) *Optional Features* (Theorem):
An optional feature α can show up on a lexical item X in the numeration only if the checking position of α is identical to the position of the first member of the X(P) chain at LF.

In the following two subsections, this result will shown to be valid for topicalization and scrambling.

6.4. Topicalization and Economy

Let us consider the consequences of the revised notions of Fewest Steps (cf. (113)) and Last Resort (cf. (114)) for licit and illicit cases of topicalization. The relevant example involving illicit *wh*-topicalization in English is repeated here once more, as (121-a).[54] Again, (121-b) is a simplified LF representation of (121-a); here, the NP *who*$_2$ has undergone LF raising from the embedded SpecTop position to the matrix SpecC position, adjoining there.

(121) a. *Who$_1$ said t$_1$ [$_{CP}$ that [$_{TopP}$ [$_{SpecTop}$ who$_2$] Top [$_{IP}$ John likes t$_2$]]] ?
b. *[$_{SpecC}$ [$_{SpecC}$ who$_1$] who$_2$] ... [$_{CP}$ [$_{TopP}$ [$_{SpecTop}$ (t'_2)] ... t$_2$]]

Recall from subsection 3.2.1 that examples like (121-a) can be ruled out by the classical Fewest Steps condition, whereas the Last Resort condition in its standard formulation is not needed in order to exclude this sentence (cf. subsection 3.4.4); recall also that the PUB excludes examples like (121-a) as involving ambiguous binding at LF

---

[53]Note incidentally that this theorem is closely related to the statement that "α enters the numeration only if it has an effect on output," which is proposed as an economy constraint in its own right in Chomsky (1995, 294).

[54]Given that reference sets are defined via LF-identity, but not via identity of numerations, what follows carries over immediately to illegitimate *wh*-topicalization in German; recall the discussion in subsection 5.2.1 above.

(subsection 4.2.6). Thus, the task now is to show that *wh*-topicalization as in (121-a) can be ruled out under the present system of economy as well. It turns out that this is indeed the case. Interestingly, however, this time both economy constraints, i.e., Fewest Steps and Last Resort, are necessary to achieve this result.

There are two possible derivations of the LF representation in (121-b) via the S-structure representation in (121-a), resulting from the fact that we are in principle free to choose or not to choose the optional [top] feature in the numeration. In derivation D_1, *who*$_2$ does not bear the feature [top] in the numeration, and neither does Top. Overt topicalization in D_1 therefore violates the Last Resort condition – recall that movement without feature checking is possible under (114), but only if that movement ends up in a typical checking position for an unchecked feature of the moved item. If *who*$_2$ does not have an unchecked [top] feature in the first place, topicalization is prohibited by Last Resort. Note, however, that the revised Fewest Steps condition (unlike the original one) does not rule out derivation D_1 – no competing derivation that yields the same LF output requires fewer instances of feature checking.

The case is different with the second derivation (D_2) of (121-b) which will be considered here. D_2 differs from D_1 in that *who*$_2$ bears the feature [top] in the numeration.[55] Hence, overt topicalization respects Last Resort. However, this time Fewest Steps is violated: There is an alternative derivation that converges and yields the same LF output but involves one instance of feature checking fewer. This derivation differs from D_2 in that a [top] feature is not present on the *wh*-phrase, so that overt *wh*-topicalization does not occur and LF raising to the matrix clause takes place from the in situ position of *who*$_2$ in the embedded clause. To sum up, in the present approach a conspiracy of Fewest Steps and Last Resort is responsible for the ban on *wh*-topicalization – there is no derivation of (121-b) via (121-a) that respects *both* economy constraints.

Note that ordinary topicalization is still permitted under these assumptions; recall (21-b) in English, which is repeated here as (122):

(122) I think [$_{CP}$ that [$_{TopP}$ Mary$_1$ John likes t_1]]

(122), in contrast to (121-a), must not be blocked by the interaction of economy constraints. And indeed, it emerges that there is a legitimate derivation of (122). In this derivation, *Mary*$_1$ is equipped with the optional [top] feature in the numeration. Hence, overt topicalization satisfies Last Resort. More interestingly, however, Fewest Steps is not violated either – there is no competing derivation that is more economical. As before, the reason is that topicalization is not undone by subsequent LF raising in (122), as it is in (121-a). Technically, the derivation generating (122) is not blocked by a

[55]Of course, this implies that Top also must have a [top] feature so as to be able to check off the [top] feature of *who*$_2$. For the sake of simplicity, I will abstract away from this issue in what follows; i.e., in all cases where I postulate that an XP has an optional feature (like [top] or [scr]) in the numeration, this feature is tacitly assumed to also be generated on a target head – otherwise, the derivation crashes.

derivation that differs minimally in having the NP object *Mary*$_1$ in situ at S-structure because the two derivations do not share identical LF outputs and hence do not compete, given the notion of reference set in (110) and given the above assumption that optional movement types like topicalization are not literally reconstructed at LF (see footnote 16).

6.5. Scrambling and Economy

Basically the same picture arises with licit and illicit cases of scrambling. Again, let us consider the illicit case of optional movement first, viz., *wh*-scrambling in German as in (123-a), with the schematic LF representation in (123-b):

(123) a. *Wie$_1$ hat [$_{IP}$ was$_2$ [$_{IP}$ der Fritz t$_1$ t$_2$ repariert]] ?
how has what$_{acc}$ ART Fritz$_{nom}$ fixed
b. *[$_{SpecC}$ [$_{SpecC}$ wie$_1$] was$_2$] ... [$_{IP}$ (t$'_2$) [$_{IP}$ ... t$_2$]]

As with *wh*-topicalization, two possible derivations that generate the S-structure representation in (123-a) and terminate in the LF representation in (123-b) must be ruled out. In the first derivation D$_1$, *was*$_2$ does not bear the feature [scr]. Hence, Last Resort is violated in D$_1$ by scrambling to IP; the Fewest Steps condition, however, is respected, because an alternative derivation with the *wh*-phrase *was*$_2$ in situ at S-structure does not involve fewer feature checking operations (again, D$_1$ would violate the original Fewest Steps condition that does not discriminate between checking and movement). In contrast, in the second derivation that is possible a priori, D$_2$, *was*$_2$ does bear the feature [scr]. Consequently, overt scrambling in front of the subject respects Last Resort; this time, though, Fewest Steps is violated: There is an alternative derivation in which *was*$_2$ does not have an optional [scr] feature and shows up in situ at S-structure. This derivation has the same LF output as D$_2$ (with *was*$_2$ adjoined to SpecC), but achieves this with one instance of feature checking fewer.

Turning next to legitimate cases of scrambling, as in (22-b), the wellformedness follows in complete analogy to that of ordinary topicalization, as in (122). (22-b) is repeated in (124):

(124) dass [$_{IP}$ den Fritz$_1$ [$_{IP}$ keiner t$_1$ gesehen hat]]
that ART Fritz$_{acc}$ no-one$_{nom}$ seen has

Suppose that (124) is generated by a derivation in which the NP *den Fritz*$_1$ is equipped with a [scr] feature. This [scr] feature is checked by left-adjunction to IP, in accordance with Last Resort. Furthermore, this derivation is not blocked by a competing derivation in the same reference set via Fewest Steps, assuming that scrambling behaves like topicalization in that it is not literally reconstructed at LF. Again, intuitively, the contrast between licit and illicit cases of scrambling is that ill-formed *wh*-scrambling as in (123-a) is later undone by LF raising, whereas well-formed ordinary scrambling as in (124) persists at LF.

To draw a preliminary conclusion, the present system of economy constraints comprising the Fewest Steps condition in (113) (together with the notion of reference set in (110)) and the Last Resort condition in (114) has the following consequences: First, it evades the conceptual objection raised in subsection 5.1 because it renders the concept of Form Chain superfluous. Second, it is immune to the empirical criticism in subsection 5.2 because it systematically permits partial *wh*-movement. And third, it accounts for all the evidence that was formerly handled by the original economy constraints (with the qualification made in footnote 48). However, it has yet to be shown how the redundancy between the PUB and economy constraints addressed in subsection 4.2.6 can be resolved. I address this issue now.

6.6. Deriving PUB Effects from Economy

In what follows, I would like to argue that typical PUB effects as they have been discussed in subsection 4.2 above follow from the interaction of the revised Fewest Steps and Last Resort conditions, and that, accordingly, the PUB can be dispensed with. In fact, it seems to me that the present system of economy constraints goes a long way towards a general theory of improper movement. I address super-raising (cf. subsection 4.2.1), long-distance scrambling (subsection 4.2.2), and the *wh*--island asymmetry between *wh*-movement and topicalization (subsection 4.2.5) in turn.

6.6.1. Super-Raising and Economy

Before I tackle super-raising, let me briefly point out the consequences that the revised economy approach developed above has for well-formed cases of successive-cyclic raising, as in (125):

(125) John$_1$ is likely [$_{IP}$ t$_1''$ to be asked [$_{IP}$ t$_1'$ to t$_1$ leave]]

Let us assume that successive-cyclic A-movement via the two intervening SpecI positions is required in (125) because of locality requirements; the constraint in question cannot be the Barriers Condition (IP is not a barrier for an item in SpecI), but it might be (something like) Relativized Minimality (Minimize Chain Links). If locality requires the presence of the traces t$_1''$ and t$_1'$ in (125), it must be ensured that their existence is compatible with economy constraints. Indeed, the present system permits successive-cyclic raising in (125): There are three movement operations to typical checking positions (for the [Case] feature) that apply to *John*$_1$, but there is only one actual checking operation (viz., in the matrix SpecI position) – *John*$_1$ is "fooled" into A-movement twice. Since in all three instances the [Case] feature of *John*$_1$ is still unchecked at the stage of the derivation at which NP-raising occurs, Last Resort is respected.[56] Moreover, Fewest Steps is respected because there is no competing derivation in the same reference set that involves fewer feature checking operations ([Case] being an obligatory feature on an NP).

Thus, in allowing successive-cyclic raising, the approach to economy developed here differs again from the standard approach: Definitions of Fewest Steps and Last Resort of the kind discussed in section 3 either require recourse to the concept of Form Chain to permit successive-cyclic raising as in (125) or necessitate the assumption that the NP $John_1$ is in fact involved in three instances of feature checking on its way to the matrix clause (cf. Chomsky (1995, 283)).

Next, consider a typical super-raising construction, as in (126):

(126) *John_1 seems that it is likely t_1 to win

As noted before, there are various derivations that need to be ruled out in order to ensure that (126) is actually predicted to be ungrammatical. A first derivation involving movement in one fell swoop violates locality; a second derivation with counter-cyclic insertion of *it* violates both locality and the Strict Cycle Condition; and a third successive-cyclic derivation via SpecC involves improper movement, which can be reduced to the PUB. The relevant derivation is repeated in (127):

(127) *John_1 seems $[_{\text{CP}}\ t'_1$ that it is likely $[_{\text{IP}}\ t_1$ to win]]

To fulfill the Barriers Condition, the NP $John_1$ moves first into the intermediate SpecC position, and then into the subject position of the matrix clause. Both operations respect strict cyclicity. What, then, rules out the combination of movement to SpecC followed by movement to SpecI if the PUB is to be dispensed with?

The answer is that (127) is excluded by economy. This time, though, it is not the *interaction* of Fewest Steps and Last Resort that achieves this; the Last Resort condition in (114) suffices to rule out (127). The NP $John_1$ bears the feature [Case], but it does not bear the feature [+wh]. Hence, by Last Resort, it may undergo movement to a typical checking position for [Case], but not to the embedded SpecC position, which is a typical checking position for [+wh].[57]

Thus, abstracting away from the qualification in the last footnote for the time being, the prohibition against super-raising is derived in toto, with the gap created by derivations like (127) that involve improper movement now being filled by an eco-

[56]Note that although the revised version of Last Resort in (114) differs from the standard version of Last Resort in (42) in that it allows movement without feature checking, both versions succeed in ruling out movement to a checking position for a [–Interpretable] feature α (like [Case]) if α has already been checked on the moved item. Thus, the instances of "unmotivated" NP-raising ruled out by Last Resort in Chomsky (1995, 284) are still ruled out by Last Resort in the present framework.

[57]A similar result is obtained in the theory of economy developed in Collins (1994a, 23-24). Like Collins' (1994a) analysis, the present approach raises the question of what happens if some item undergoes successive-cyclic super-raising that *does* bear a [+wh] feature. I address this question in subsection 6.7.

nomy constraint (Last Resort) instead of a cyclicity constraint (the PUB).

6.6.2. Long-Distance Scrambling and Economy

Next, consider the prohibition against long-distance scrambling from finite clauses in German, as in (128):

(128) *dass Antje Hygrometer$_1$ sagt [$_{CP}$ t$'_1$ dass niemand t$_1$ mag]
that Antje hygrometers says that no-one likes

If movement does not apply successive-cyclically, via the embedded SpecC position, the Barriers Condition is violated (cf. subsection 4.2.4). But what rules out long-distance scrambling via SpecC, given that we no longer want to resort to an account in terms of the PUB? Not surprisingly, successive-cyclic long-distance scrambling as in (128) is ruled out by economy, in more or less the same way as successive-cyclic super-raising (cf. (126)): The NP *Hygrometer*$_1$ in (128) has a [scr] feature, but no [+wh] feature. Hence, intermediate substitution in SpecC violates Last Resort – this position is a typical checking position for the feature [+wh] but not for the feature [scr]. Thus, the PUB is no longer needed to derive the prohibition against long-distance scrambling in German.[58]

6.6.3. Wh-Island Asymmetries and Economy

The third case of ambiguous binding encountered above concerns the asymmetry between *wh*-movement and topicalization in the case of *wh*-islands in German. The relevant contrast was that between (91-a), which has an intermediate (Subjacency-like) status, on the one hand and (92-a), which typically strikes speakers as considerably worse, on the other. These examples are repeated here:

(129) a. ??[$_{CP}$ – C [$_{TopP}$ Radios$_1$ weiß ich nicht [$_{CP}$ wie$_2$ [$_{TopP}$ t$'_1$ man t$_2$ t$_1$
radios$_{acc}$ know I not how one
repariert]]]]
fixes

b. ?*[$_{CP}$ Was$_1$ C [$_{TopP}$ – weißt du nicht [$_{CP}$ wie$_2$ [$_{TopP}$ (t$'_1$) man t$_2$ t$_1$
what$_{acc}$ know you not how one
repariert]]]] ?
fixes

[58] A potential problem is that long-distance scrambling from finite clauses is in fact permitted in a number of languages, among them Russian (with arguments and adjuncts), Korean, and Japanese (in these two languages arguments can undergo long-distance scrambling, but adjuncts cannot; see above). This issue is addressed in Müller & Sternefeld (1993); the proposal there is independent of the PUB and could be carried over into the present approach basically unchanged.

Suppose, following the reasoning in subsection 4.2.5 above, that the key to a solution of this problem is that an intermediate trace in the embedded SpecTop position is possible with long-distance topicalization, as in (130-a), but not with long-distance *wh*-movement, as in (130-b).

(130) a. ??[$_{CP}$ – C [$_{TopP}$ Radios$_1$ ... [$_{CP}$ wie$_2$ [$_{TopP}$ t_1' man t_2 t_1 repariert]]]
b. ?*[$_{CP}$ Was$_1$ C [$_{TopP}$ – ... [$_{CP}$ wie$_2$ [$_{TopP}$ t_1' man t_2 t_1 repariert]]]] ?

As shown above, the illegitimacy of the intermediate trace t_1' in SpecTop in (130-b) follows from the PUB. But, again, the same result is obtained in the revised economy approach without invoking an explicit constraint against improper movement: *Wh*-movement via SpecTop is ruled out by the interaction of Fewest Steps and Last Resort. A derivation of (130-b) in which *was*$_1$ does not bear the optional [top] feature in the numeration (in addition to its obligatory [+wh] feature) does not violate Fewest Steps, but it does violate Last Resort. In contrast, a minimally different derivation in which *was*$_1$ initially has a [top] feature respects the Last Resort condition. However, this derivation is ruled out by Fewest Steps because there is an alternative derivation that converges and yields the same LF output (i.e., that is in the same reference set) but involves one instance of feature checking fewer; this derivation involves *wh*-movement across the *wh*-island.[59]

More generally, then, it seems reasonable to venture the hypothesis that the interaction of the economy constraints Fewest Steps and Last Resort in their formulations in (113) and (114) suffices to account for illegitimate combinations of movement operations and renders specific constraints on improper movement (like the PUB) superfluous.[60]

---

[59]This reasoning presupposes that a non-V/2 Top head may in principle check a [top] feature in German (and that the obligatory filling of Top by V/2 movement in topicalization contexts results from actual lexical filling of SpecTop at S-structure); if this were not the case, the Fewest Steps condition would in fact not be needed to account for the illformedness of (130-b). For lack of decisive evidence, I will remain uncommitted as far as this issue is concerned; recall the remarks in subsection 3.4.2 above.

[60]Note, however, that there are still certain empirical differences between a PUB-based account and the revised economy approach to improper movement. One will be discussed in the following subsection; another one concerns illicit sentences such as (i) in English:

(i) *Who$_1$ t_1 thinks [$_{CP}$ what$_2$ John saw t_2] ?

Epstein (1992) argues that overt *wh*-movement to a [–wh] position (the embedded SpecC position in (i)) followed by LF raising to a [+wh] position (the matrix SpecC position) is precluded by economy. Similarly, Müller & Sternefeld (1996) show that the illformedness of (i) follows from the PUB (overt substitution in SpecC is followed by LF adjunction to SpecC). In the present approach, however, Fewest Steps and Last Resort permit sentences like (i) just as they permit partial *wh*-movement in German. Accordingly, I would like to contend that whatever is responsible for the lack of partial *wh*-movement in English also covers (i); cf. the references given in footnote 52. Consequently, unless further assumptions are made, examples like (i) are predicted to be well formed in German under the revised economy approach. The data appear not to be crystal-clear. Whereas an ungrammatical German example of the type in (i) is given in Müller

6.7. A Problem and Its Solution

Finally, I would like to address a problem that arises in the present approach. This problem concerns the interaction of the two movement types raising and *wh*-movement and has already been alluded to in footnote 57. Given that both [Case] and [+wh] are obligatory features (in the sense that they must show up on a *wh*-NP), nothing rules out a combination of raising and *wh*-movement as in (131), which is of course a welcome result.

(131) $[_{CP}$ Who$_1$ $[_{IP}$ t_1' was kissed t_1]] ?

Both movements respect Last Resort, and the derivation is not blocked by Fewest Steps either because there is no competing derivation with fewer instances of feature checking (if one or both of the obligatory features [Case] and [+wh] are not checked, non-convergence results).

The problem now is that the same considerations lead us to conclude that successive-cyclic derivations of super-raising constructions as in (132) are also well formed, contrary to fact ((132) differs minimally from the typical super-raising example (126) discussed above in that the [–wh] NP *John* is replaced by the [+wh] NP *who*).

(132) *Who$_1$ t_1'' seems $[_{CP}$ t_1' that it is likely $[_{IP}$ t_1 to win]] ?

This derivation does not violate Fewest Steps because it involves only the minimal number of feature checking operations required for convergence – the [Case] feature of *who*$_1$ is checked in the matrix SpecI position, and the [+wh] feature of *who*$_1$ is checked in the matrix SpecC position. (132) does not violate Last Resort either: The first movement operation to the position of t_1' ends up in a typical checking position (SpecC) for an unchecked feature ([+wh]) of *who*$_1$, and the next two movement operations result in actual feature checking. Thus, it appears that the account of super-raising is undermined if the super-raised NP bears a [+wh] feature.

Note that in contrast to the revised economy approach, the PUB rules out (132)–A-bar movement to SpecC is illegitimately followed by raising to another kind of position, viz., SpecI. On the other hand, the PUB permits (131) only because it is explicitly defined as a constraint on variables (conceived of as locally A-bar bound traces); there *is* ambiguous binding of a trace in (131), but that trace (t_1) is not a variable and hence is immune to PUB effects, by stipulation (recall the discussion in subsection 4.2.3). I think this can be taken to indicate that a departure from the "best case" is necessary in any event and that an additional stipulation in the revised economy approach to rule out derivations like (132) does not imply that the analysis is inferior to the

---

& Sternefeld (1996, 485), others (e.g., McDaniel (1989), d'Avis (1993, 92-95), Brandner (1994), Brody (1995, 108), and Fanselow & Mahajan (1996, 10)) assume that well-formed sentences of this kind do in fact exist in (colloquial) German. I have to leave this issue undecided.

PUB approach. In other words, there is a trade-off between the PUB-based account of improper movement and the economy-based account argued for in this section: Whereas the PUB requires an additional stipulation to permit (131), economy requires an additional stipulation to exclude (132).

How, then, should the economy-based approach be revised? Various options come to mind; I will here follow proposals that go back to Williams (1974) (cf. the so-called "Williams cycle") and Sternefeld (1992) and pursue the idea that the Last Resort condition is to be modified in such a way that it predicts a certain sequence of feature checking.[61] For the sake of concreteness, suppose that the set of features on a lexical item is ordered and that this order is not arbitrary, but rather reflects the hierarchy of typical checking positions in the tree, as in (133):[62]

(133) [+wh] $\succ$ [top] $\succ$ [scr] $\succ$ [Case]

Based on (133), I would like to suggest that movement to a typical checking position for a feature γ is possible only if there is no unchecked lower-ranked feature δ on the lexical item. This can be accomplished by revising the Last Resort condition (114) as in (134):

(134) *Last Resort* (second revision):
α is raised to a position β only if β is a typical checking position for the *lowest-ranked* unchecked morphological feature of α.

This revision leaves all previous reasonings intact.[63] In addition, however, (134) rules out *wh*-super-raising as in (132) while at the same time allowing (131). Thus, (131) does not violate Last Resort in (134) because the [Case] feature (which is lower than the [+wh] feature on the hierarchy in (133)) is checked off before the [+wh] feature is checked off, but (132) is now ruled out by Last Resort: The embedded SpecC position, albeit a typical checking position for an unchecked feature of *who*$_1$ (viz., [+wh]) at the point in the derivation where the first movement operation applies to this item, is nevertheless not a typical checking position for its lowest-ranked unchecked feature, which is not yet [+wh], but still [Case]. More generally, Last Resort now predicts that the [Case] feature must be checked off by an NP before other movement operations

[61] For alternative approaches to the problem of how to exclude super-raising of *wh*-items as in (132) by economy constraints, cf. Collins (1994a, 24) and Costa (1995, 6f).

[62] See Chomsky (1993, 28) for similar suggestions with respect to verbal features, made in order to derive the effects of the Mirror Principle (cf. Baker (1988)) in a theory where inflection is base-generated on the verb.

[63] In particular, all the cases where the contribution of the Fewest Steps condition was indispensable to derive ungrammaticality continue to rely on this transderivational economy constraint. The reason is that the illicit combinations of movement operations discussed so far (except for super-raising and long-distance scrambling) can still fulfill Last Resort as in (134); and super-raising and long-distance scrambling could be ruled out by Last Resort alone already under the earlier formulation in (114).

can apply; this result seems to be correct.

7. Conclusion

To sum up, a derivational theory of movement has emerged in the preceding sections that crucially incorporates the notions of locality, economy, and cyclicity. As far as locality is concerned, I have adopted the Barriers Condition (see (1)). With respect to cyclicity constraints (conceived of as constraints that govern the interaction of movement operations), I have argued that in addition to the Strict Cycle Condition (see (74)), a constraint banning improper movement seems necessary under standard notions of economy; I have suggested that this role is played by the Principle of Unambiguous Binding (PUB) (see (82)). However, it turned out that the interaction of independently motivated revisions of the economy constraints Fewest Steps (see (113)) and Last Resort (see (134)) comes close to a principled theory of improper movement and thereby renders a constraint like the PUB superfluous.

With this theory of movement as the background, I will now return to the issue of remnant movement and address the four questions posed at the end of chapter 1. The guiding idea will be that the derivational approach to movement developed in this chapter significantly contributes to an explanation of the properties of incomplete category fronting.

Chapter 3

Proper Binding

1. The Problem

Recall a typical example involving incomplete category fronting in German, such as (1):

(1) $[_{VP}$ t_2 Gelesen $]_1$ hat $[_{IP}$ $[_{NP}$ das Buch $]_2$ $[_{IP}$ keiner t_1 $]]$
read has the book$_{acc}$ no-one$_{nom}$

Under the remnant movement approach adopted in chapter 1, (1) is derived by two movement operations: NP_2 is scrambled out of VP_1, and VP_1 is topicalized. This latter process creates an unbound trace, t_2. The problem now is that constructions involving unbound traces do not otherwise appear to be legitimate.[1] Thus, unbound traces created by *wh*-lowering regularly induce illformedness, as is shown by the following English example (= (47) of chapter 1):

(2) *John asked t_1 $[_{CP}$ who$_1$ Mary saw Bill]

As noted in subsection 3.1 of chapter 1, there does not seem to be anything wrong with selection requirements in (2). The matrix verb *ask* selects an embedded *wh*-clause, and this implies (in English) that either SpecC or C must be lexicalized by an appropriate lexical item carrying a [+wh] feature. The lowered *wh*-phrase *who*$_1$ is a priori a good candidate for satisfying this requirement. Thus, it looks as though it is the unbound trace of *who*$_1$ that is responsible for the ungrammaticality of (2).

Similarly, as noted in subsection 3.1 of chapter 1, topic-lowering as in (3) in German (= (48) of chapter 1) is impossible, and it seems plausible to relate the impossibility of (3) to the occurrence of an unbound trace.

[1] This is often considered to be one of the most problematic aspects of the remnant movement approach to incomplete category fronting. See, e.g., Haider (1991, 28).

(3) *dass Fritz t_1 sagte [$_{CP}$ dem Peter$_1$ habe [$_{IP}$ keiner die Claudia
that Fritz said ART Peter$_{dat}$ has$_{subj}$ no-one$_{nom}$ ART Claudia$_{acc}$
gesehen]]
seen

In what follows, I would like to discuss a number of proposals that have been made to account for illicit lowering operations as in (2) and (3) and see whether or not they can be reconciled with the very existence of unbound traces in remnant movement configurations.

2. The PBC at S-Structure

Fiengo (1977) discussed ill-formed English sentences like the following:

(4) a. *t_1 was destroyed [$_{NP}$ the city]$_2$ by [$_{NP}$ the barbarians]$_1$
b. *t_1 was [$_{NP}$ a fly]$_1$ on the wall

In line with standard assumptions of the time, he assumed that well-formed English passive sentences such as (5-a) and existential sentences such as (5-b) are derived by two operations. First, NP Postposing takes place (removing the subject NP from its canonical pre-verbal position and locating it farther to the right), and second, the subject position is filled by the object NP in the passive construction and by the expletive *there* in the existential construction:

(5) a. [$_{NP}$ The city]$_2$ was destroyed t_2 by [$_{NP}$ the barbarians]$_1$
b. There was [$_{NP}$ a fly]$_1$ on the wall

According to this view, the illicit sentences in (4) correspond to intermediate stages in the derivations of the sentences in (5). The question that arose in the framework adopted by Fiengo was why NP Postposing as in (4) could not apply unless subsequent operations that move or insert another item (an NP or an expletive) into the vacated subject position took place.

For similar reasons, the same problem arose with illicit NPs like the one in (6-a). Given that (6-b) is derived by NP Postposing followed by NP Preposing, it is clear that (6-a) is an intermediate stage in the derivation of (6-b), and the question is why this intermediate stage does not qualify per se as a well-formed grammatical object.

(6) a. *t_1 destruction (of) [$_{NP}$ the city]$_2$ by [$_{NP}$ the barbarians]$_1$
b. [$_{NP}$ The city's]$_2$ destruction t_2 by [$_{NP}$ the barbarians]$_1$

Fiengo's idea was that the illformedness of the examples in (4) and (6-a) can be traced back to the fact that the initial subject trace t_1 occurs unbound at S-structure. To exclude such unbound traces, he postulated the Proper Binding Condition (PBC) as an S-structure filter. The PBC can be formulated as follows:

(7) *The Proper Binding Condition* (PBC):
Traces must be bound at S-structure.

Assuming the standard notion of binding in terms of c-command (cf., e.g., Chomsky (1981)), the traces t_1 in (4) and (6-a) violate the PBC because they are not c-commanded by their antecedents, which appear in lower positions at S-structure.[2]

It seems that the original motivation for the PBC as an S-structure filter has largely been lost over the last 20 years. On the one hand, the sentences in (4) are standardly taken to violate a constraint like the Extended Projection Principle (EPP) (see Chomsky (1982; 1995)) that requires the subject position to be filled at S-structure; similar considerations apply in the case of the pre-noun position in English NPs (cf. (6-a)). And on the other hand, the assumption that NP preposing or *there*-insertion may overwrite traces in θ-positions does not seem to be tenable in current theorizing anyway. Nevertheless, Fiengo's (1977) PBC in (7) directly rules out *wh*-lowering and topic-lowering, as in (2) and (3), and thus receives independent corroboration. Unfortunately, it is evident that remnant movement structures as in (1) also violate the PBC at S-structure and are therefore incorrectly predicted to be ungrammatical. For this reason, I take it that the PBC as an S-structure filter is not correct and that an alternative explanation of the illegitimacy of syntactic lowering must be found.

3. The Generalized PBC

Lasnik & Saito (1992, 90) propose a derivational formulation of the PBC. Empirical motivation for this can be gained from ungrammatical examples like (8), which exemplifies a typical complementizer-trace effect in English:[3]

(8) *Who$_1$ do you think [$_{CP}$ t_1' that [$_{IP}$ t_1 left early]]?

Suppose for the sake of the argument that complementizer-trace effects do not result from a failure of some form of head-government (as in Rizzi (1990), Chomsky & Lasnik (1993), and much related work), but rather from a violation of the Barriers

[2] Cf. May (1977). Note, incidentally, that Fiengo (1977, 45) assumes that binding is not defined via c-command (as has become the standard view), but rather via precedence: A trace must be preceded by its antecedent. Clearly, as it stands this rules out all rightward movement. Concerning Heavy NP Shift (HNPS), Fiengo (1977, 48) argues that this result is tenable. He suggests that HNPS structures are in fact derived by leftward movement, a process very much akin to Larson's (1988; 1991) idea of Light Predicate Raising; also see Kayne (1994). As far as other rightward movement operations are concerned, they are either ignored in Fiengo's analysis (e.g., this is true of extraposition), or they are permitted by an explicit stipulation that the PBC only affects NP traces (as in the case of Particle Shift and Affix Hopping).

[3] Actually, the example considered by Lasnik & Saito (1992) is somewhat different; it involves *wh*-extraction from a subject NP. However, their argument for the reformulation of the PBC based on subject NP islands rests on a problem that does not arise in the theory of barriers adopted here; cf. section 2 of chapter 2. The example considered here makes the same point, though.

Condition, essentially as proposed in Chomsky (1986, 47-48). On this view, the presence of a complementizer *that* in English induces a barrier.[4] If the derivation proceeds as depicted in (8), the illformedness does not pose a problem – movement from t_1 to t'_1 violates the Barriers Condition, and given that t_1 is thereby marked * (cf. subsection 3.2.5 of chapter 2) and cannot be deleted on the way to LF (in contrast to t'_1, which is deleted under Chomsky & Lasnik's (1993) assumptions about uniform chains), strong ungrammaticality is indeed to be expected.

However, as noted by Barss (1985) and, following him, Chomsky (1986, 87), there is a potential loophole for movement in (8)– what if the subject is first lowered from the SpecI position to a VP-adjoined position and then raised again via the embedded SpecC position to the matrix SpecC position, as in (9)?

(9) *Who$_1$ do you think [$_{CP}$ t''_1 that [$_{IP}$ t_1 [$_{VP}$ t'_1 [$_{VP}$ left early]]]]?

Crucially, the first step here, i.e., lowering from SpecI to an adjoined position in the extended projection of the verb, does not violate the Barriers Condition. Of course, the second movement operation, which moves the subject from the adjoined position to the SpecC position, does violate the Barriers Condition under present assumptions. On the one hand, however, an intermediate, deletable argument trace t'_1 marked * cannot induce as strong a degree of ungrammaticality as is typical with complementizer-trace effects; on the other hand, we would expect whatever proviso is made in the case of adjunct and object movement across a lexical complementizer in English (cf. the last footnote) to carry over immediately to (9). Indeed, if (9) were an admissible derivation, we would wrongly expect (8) to be on a par with examples like (10) (cf. Lasnik & Saito (1984, 256; 1992, 68)), in which a complementizer-trace effect does not show up because the initial trace t_1 is not adjacent to a lexical complementizer and is thus "rescued" by an intermediate trace t'_1 (which in turn is immune to the complementizer-trace environment):

(10) Who$_1$ do you believe [$_{CP}$ t''_1 that Mary said [$_{CP}$ t'_1 [$_C$ –] [$_{IP}$ t_1 left early]]] ?

Thus, it seems that lowering as in (9) must be excluded in order to maintain the derivation of complementizer-trace effects in English. A standard PBC along the lines of Fiengo (1977) does not achieve this result, though. The reason is that the PBC is formulated there as an S-structure filter; but although there is a step in the derivation at which a trace occurs unbound in (9) – t_1 is unbound after *wh*-lowering to VP (step

[4]Two remarks are in order here. First, whether the barrier created by a lexical complementizer in English is C', as argued in Chomsky (1986), or IP, as argued in Koster (1987, ch.4) and Müller (1995, 31-38), is immaterial for present purposes. (However, IP barrierhood would follow directly under the assumptions in section 2 of chapter 2 concerning the notion of barrier, given that I and C are in fact distinct when C is filled by a complementizer.) And second, something must be said about why object and adjunct movement can freely cross the purported barrier induced by the complementizer, despite an apparent violation of the Barriers Condition. See Müller (1995, ch. 2) for a discussion of both issues.

(1)) and before *wh*-raising to SpecC (step (2)) –, every trace of *who*$_1$ is bound at S-structure, and the PBC is fulfilled. This problem disappears under Lasnik & Saito's (1992, 90) reformulation of the PBC as the Generalized Proper Binding Condition (GPBC):

(11) *Generalized Proper Binding Condition* (GPBC):
Traces must be bound throughout a derivation.

The GPBC not only excludes lowering in (9) but also accounts for the illformedness of the lowering derivations in (2) and (3).[5] However, since the GPBC is a *strengthened* version of the PBC, it clearly does not fare any better than the original PBC with respect to cases of remnant movement as in (1). These constructions are still wrongly classified as ungrammatical.

Furthermore, it is worth pointing out that quite independently of the lowering problem, the derivation in (9) is ruled out by the interaction of the economy constraints Fewest Steps and Last Resort, as a PUB effect (cf. section 6 of the previous chapter). To see this, consider the movement operations that take place in (9) (which is repeated here) more closely:

(12) *Who$_1$ do you think [$_{CP}$ t$_1''$ that [$_{IP}$ t$_1$ [$_{VP}$ t$_1'$ [$_{VP}$ left early]]]]?

The first step is adjunction to VP, i.e., scrambling, according to the assumptions of chapter 2. The second and third steps, then, involve substitution in SpecC, i.e., *wh*-movement. There are two 'sub'derivations of (12) to be examined. In one derivation (D$_1$), the *wh*-phrase *who*$_1$ is equipped with a [+wh] feature in the numeration but not with a [scr] feature. Hence, intermediate scrambling violates Last Resort (cf. (134) from chapter 2, repeated here).

(13) *Last Resort*:
α is raised to a position β only if β is a typical checking position for the lowest-ranked unchecked morphological feature of α.

The Fewest Steps condition (cf. (113) from chapter 2, which is also repeated here) is not violated by D$_1$, though – there is no competing derivation in the same reference set (i.e., converging and yielding the same LF output) that involves fewer instances of feature checking.

(14) *Fewest Steps*:
If two derivations D$_1$ and D$_2$ are in the same reference set and D$_1$ involves fewer checking operations than D$_2$, then D$_1$ is to be preferred over D$_2$.

[5]Note also that the GPBC is incompatible with the kinds of derivations envisaged by Fiengo (1977) that were discussed in the previous subsection, because these derivations depended on the possibility that unbound traces in the derivation could be prevented from inducing ungrammaticality by later operations which apply before S-structure and remove such traces.

The case is different with the other derivation of (12) that is a priori possible, viz., D_2. In this derivation, *who* bears a [scr] feature in the numeration, so that neither intermediate scrambling nor the following two instances of *wh*-movement violate Last Resort.[6] However, D_2 violates Fewest Steps because there is an alternative derivation that is more economical in the sense that it dispenses with the optional [scr] feature and hence with the additional [scr] feature checking operation. Thus, it emerges that the impossibility of a derivation like (12) is explained straightforwardly as an instance of improper movement – it involves a PUB effect, which can be derived from the interaction of Fewest Steps and Last Resort.[7]

Although we now have additional corroboration of the approach to movement theory developed in chapter 2, it seems that we are not yet closer to the original goal of this chapter, which was to come up with a constraint on lowering (as in (2) and (3)) that does not simultaneously exclude remnant movement. To this end, I will consider another version of the PBC in the following subsection.

4. The PBC at LF

4.1. The Approach

In line with the view embodied in the minimalist program that S-structure does not exist as a level of representation to which constraints may apply, Truckenbrodt (1992) suggests that the Proper Binding Condition should be taken to hold only at the interface level LF, as in (15).

(15) *The Proper Binding Condition* (revised):
Traces must be bound at LF.

The basic idea pursued by Truckenbrodt is then that topicalized items undergo reconstruction at LF for semantic reasons. Among other things, this implies that a sentence involving remnant topicalization, such as (1), repeated here as (16-a), ends up with an LF representation like (16-b), where the topicalized VP is reconstructed into its base position and the VP-internal scrambling trace t_2 is bound, in accordance with the PBC

---

[6]Recall that the [scr] feature is ranked lower than the [+wh] feature (cf. subsection 6.7 of chapter 2) and is accordingly checked off first, in line with Last Resort, and that the first *wh*-movement operation in (12) is permitted by Last Resort even though it does not involve actual feature checking.

[7]Note that the derivation in (12) proceeds via what has been called Yo-Yo movement and closely resembles the illicit derivation (38) of the ungrammatical Ewe example (37-b) discussed in subsection 3.3.1 of chapter 2. Recall that the transderivational Shortest Path Condition can be invoked to rule out that latter derivation (cf. Collins (1994)). Hence, it is not surprising that the Shortest Path Condition also excludes (12), just as the interaction of Fewest Steps and Last Resort has been shown to do. I will come back to this issue shortly.

in (15):[8]

(16) a. [$_{\text{VP}}$ t$_2$ Gelesen]$_1$ hat [$_{\text{IP}}$ [$_{\text{NP}}$ das Buch]$_2$ [$_{\text{IP}}$ keiner t$_1$]]
read has the book$_{acc}$ no-one$_{nom}$

b. – hat [$_{\text{IP}}$ [$_{\text{NP}}$ das Buch]$_2$ [$_{\text{IP}}$ keiner [$_{\text{VP}}$ t$_2$ gelesen]$_1$]]

Thus, it looks as though remnant movement structures are permitted by assuming that the PBC holds only at LF. The question now is, of course, whether illicit S-structure lowering is still ruled out. If we consider the failed attempt at question formation in English in (2) (repeated here), it seems that it is:

(17) *John asked t$_1$ [$_{\text{CP}}$ who$_1$ Mary saw Bill]

There are two possibilities to be considered. One is that the LF representation of (17) corresponds roughly to the S-structure representation, i.e., no further movement operation applies to the *wh*-phrase *who*$_1$. In that case, the PBC in (15) is violated at LF, and illformedness results, as desired. The other possibility is that the *wh*-phrase raises back into the matrix clause at LF, substituting in the matrix SpecC position, as shown in (18):

(18) who$_1$ John asked t$_1$ [$_{\text{CP}}$ t$'_1$ C$_{[}$+wh] Mary saw Bill]

Thus, if (17) could result in an LF representation like (18) without violating a principle of grammar, it would mean something like (19):

(19) Who$_1$ did John ask t$_1$ [$_{\text{CP}}$ whether Mary saw Bill] ?

However, although the derivation in (18) does not violate the PBC at LF, it does violate other principles. Specifically, as noted above, it is a fact about *wh*-clauses in English that they require lexicalization of either SpecC or C by an appropriate *wh*-item at S-structure (the [+wh] feature of C is strong in this language), and this requirement is of course violated by the matrix clause in (17) if it is to be interpreted as in (18)/(19). Since similar considerations apply in the case of the illicit German example involving downward topicalization in (3), it seems at first sight that the PBC in (15) successfully accounts for both the legitimacy of remnant movement and the illegitimacy of syntactic lowering. However, closer scrutiny reveals that this conclusion might be premature.

---

[8]To simplify matters, I here ignore other LF operations that arguably take place in the derivation of (16-b) from (16-a), such as quantifier raising applying to *keiner* ('no-one') and V/2 reconstruction; these operations do not affect the issue under discussion. Also, for the time being I would like to abstract away from the question of whether (VP) reconstruction is handled in terms of downward movement at LF (as argued by Chomsky (1977) and May (1985)) or by invoking the copy theory of movement (cf. Chomky (1981; 1993)); Truckenbrodt (1992) adopts the latter approach. Clearly, if an LF lowering approach were to be adopted, this would imply that LF lowering does not leave a trace, given the PBC in (15). See below.

4.2. Problems

Clearly, in order for the solution just envisaged to work, it must be ensured that all legitimate instances of remnant movement can be reconstructed and that no instance of illegitimate syntactic lowering can be undone by subsequent raising. In what follows, I show that both assumptions are problematic.

4.2.1. Syntactic Lowering, Undone by Raising

Let me begin by pointing out two cases of syntactic lowering that do not violate the PBC at LF but are nevertheless prohibited. The examples to be discussed differ from the English case in (17) in that a derivation like (18) does not appear to be excluded a priori.

Partial Wh-Movement in German The first case involves partial *wh*-movement in German. Recall that in this kind of construction, a *wh*-phrase undergoes syntactic movement to a [–wh] SpecC position, whereas its ultimate [+wh] target SpecC position is filled by a scope marker *was*. If there are [–wh] SpecC positions intervening between the moved *wh*-phrase and its target position, these positions are either filled by scope markers too or left empty, in which case C is filled by the complementizer *dass* ('that'). Relevant examples that show this are given in (20) (cf. the discussion in subsection 5.2.2 of chapter 2):[9]

(20) a. Was$_1$ meinst du [$_{CP_4}$ was$_1$ C er gesagt hat [$_{CP_5}$ wen$_1$ er t$_1$
[+wh] think you [+wh] he said has whom$_{acc}$ he
getroffen hat]] ?
met has

b. Was$_1$ meinst du [$_{CP_4}$ – dass er gesagt hat [$_{CP_5}$ wen$_1$ er t$_1$ getroffen
[+wh] think you that he said has whom$_{acc}$ he met
hat]] ?
has

As noted before, the existence of partial *wh*-movement in German shows that the strong [+wh] feature of a C node can be checked by insertion of a scope marker *was*. Furthermore, for reasons which I will not discuss here, a *wh*-phrase is required to move to a SpecC position in this construction (in apparent violation of Procrastinate)

---

[9] Note that there is no general agreement as to whether sentences like (20-b) are completely grammatical; van Riemsdijk (1982), McDaniel (1989, 576), and Haider (1993, 98), e.g., assume that they are not. I will here follow the judgement in Höhle (1990, 2; 1996, 40) and Müller & Sternefeld (1996, 507-509) according to which (20-b) is possible (also cf. Stechow & Sternefeld (1988, 358) for remarks on this issue). However, the wellformedness of (20-b) is not crucial for the argument that follows; (20-a) would suffice to make the point.

even though its scope is indicated by the scope marker:

(21) a. *Was$_1$ glaubst du [$_{CP_4}$ was$_1$ er gesagt hat [$_{CP_5}$ was$_1$ er wen$_1$
[+wh] believe you [+wh] he said has [+wh] he whom$_{acc}$
getroffen hat]] ?
met has

b. *Was$_1$ glaubst du [$_{CP_4}$ was$_1$ er gesagt hat [$_{CP_5}$ dass er wen$_1$
[+wh] believe you [+wh] he said has that he whom$_{acc}$
getroffen hat]] ?
met has

Interestingly, the principles of grammar do not determine exactly which SpecC position the *wh*-phrase undergoes overt movement to – any SpecC position between the clause the *wh*-phrase originates in and the clause that contains its eventual target position is fine:

(22) a. Was$_1$ glaubst du [$_{CP_4}$ wen$_1$ er gesagt hat [$_{CP_5}$ t$_1''$ dass sie meint
[+wh] believe you whom$_{acc}$ he said has that she thinks
[$_{CP_6}$ t$_1'$ dass sie t$_1$ liebt]]] ?
that she loves

b. Was$_1$ glaubst du [$_{CP_4}$ was$_1$ er gesagt hat [$_{CP_5}$ wen$_1$ (dass) sie
[+wh] believe you [+wh] he said has whom$_{acc}$ that she
meint [$_{CP_6}$ t$_1'$ dass sie t$_1$ liebt]]] ?
thinks that she loves

c. Was$_1$ glaubst du [$_{CP_4}$ was$_1$ er gesagt hat [$_{CP_5}$ was$_1$ sie meint [$_{CP_6}$
[+wh] believe you [+wh] he said has [+wh] she thinks
wen$_1$ (dass) sie t$_1$ liebt]]] ?
whom$_{acc}$ that she loves

With this in mind, consider now the following ungrammatical sentences, which are minimal variants of those in (20-a) and (20-b):

(23) a. *Was$_1$ meinst du [$_{CP_4}$ was$_1$ C er t$_1$ gesagt hat [$_{CP_5}$ wem$_1$ er
[+wh] think you [+wh] he said has whom$_{dat}$ he
geschlafen hat]] ?
slept has

b. *Was$_1$ meinst du [$_{CP_4}$ dass er t$_1$ gesagt hat [$_{CP_5}$ wem$_1$ er geschlafen
[+wh] think you that he said has whom$_{dat}$ he slept
hat]] ?
has

The difference between (20-a) and (20-b) on the one hand and (23-a) and (23-b) on the other is that the *wh*-phrase that ends up in the SpecC position of CP$_5$ at S-structure is base-generated within CP$_5$ in the former examples, and within the embedding clause CP$_4$ in the latter ones. However, it seems that all typical constraints on *wh*-scope marking structures are respected in (23-a) and (23-b) – the *wh*-phrase is in SpecC at S-structure, and the SpecC positions intervening between the *wh*-phrase and its LF target position are either filled by scope markers or left empty. Furthermore, and this is relevant in the present context, it is not quite clear why the PBC should not be fulfilled at LF in these examples via LF raising of the *wh*-phrase, as in (24):

(24) wem$_1$ meinst du [$_{\mathrm{CP}_4}$ t$_1''$ C er t$_1$ gesagt hat [$_{\mathrm{CP}_5}$ t$_1'$ er geschlafen hat]] ?

If the principle that precludes overt lowering is the PBC in (15), there is no obvious reason why the sentences in (23) should not co-exist alongside the grammatical version involving successive-cyclic long-distance movement in (25). It is also worth noting that the economy constraints Fewest Steps and Last Resort would not be violated by a derivation that creates the LF representation (24) from the S-structure representations in (23-a) and (23-b). (On this and the Shortest Path Condition, see below.)

(25) Wem$_1$ meinst du [$_{\mathrm{CP}_4}$ t$_1'$ dass er t$_1$ gesagt hat [$_{\mathrm{CP}_5}$ dass er geschlafen
whom$_{dat}$ think you that he said has that he slept
hat]] ?
has

Thus, illicit lowering in partial *wh*-movement constructions as in (23-a) and (23-b) is not explained under the assumption that the PBC holds only at LF, as in (15).

Yo-Yo Movement in Ewe More or less the same holds for illicit Yo-Yo movement constructions in Ewe, as discussed in subsection 3.3.1 of chapter 2. The crucial example (= (37-b) from chapter 2) is repeated here:

(26) Kofi$_1$ e me gble na t$_1$ [$_{\mathrm{CP}}$ be *wo/é fo Kosi]
Kofi FOC I said to that he hit Kosi
‘It was Kofi that I told that he hit Kosi.’

Recall that the embedded subject pronoun ‘he’ can be changed from the regular form *é* to *wo* only if the adjacent SpecC position is filled. What must be ensured, then, is that a derivation as in (27) is precluded. Here, S-structure lowering to SpecC triggers morphological alternation in the embedded SpecC position, and subsequent raising to the matrix SpecC position guarantees that all traces are bound.[10]

[10]As it stands, this analysis depends on the assumption that focus movement in Ewe is formally identical to *wh*-movement, in that it targets SpecC positions (cf. Collins (1993; 1994)). However, the consequences would be basically the same under the assumption that focus movement in Ewe targets a different landing site. Also, recall from footnote 21 of chapter 2 that Collins (1994, 54) assumes that focus-lowering in (27) is preceded by VP-adjunction. Clearly, under present assumptions such a combination of movement types would automatically qualify as a PUB effect, reducible to the interaction of Fewest Steps and Last Resort. Hence, such a derivation would not pose a problem in the first place, irrespective of the issue of lowering.

(27) *Kofi$_1$ e me gble na t$_1$ [$_{CP}$ t$'_1$ be wo fo Kosi]

Evidently, the only significant difference between illicit Yo-Yo movement in Ewe and illicit partial *wh*-lowering in German as in (23-a) and (23-b) is that the sequence *lowering to SpecC ≻ raising to SpecC* applies in its entirety before Spell-Out in Ewe but is split up by the Spell-Out operation in German. Hence, it does not come as a surprise that the PBC at LF cannot rule out Yo-Yo movement in Ewe either – at LF the "offending" trace created by intermediate lowering (viz., t$_1$) is bound. As before, the derivation in (27) is also not excluded by the interaction of Fewest Steps and Last Resort, as presented in chapter 2 (again, see below on how the Shortest Path Condition relates to all this).

Thus, I conclude that assuming the PBC to hold at LF does not suffice to systematically preclude illicit lowering. What is more, the next subsection shows that it does not even permit all well-formed cases of remnant movement.

4.2.2. Remnant Movement That Is Not Reconstructed

Recall that the assumption that traces must be bound at LF can be reconciled with the existence of remnant VP topicalization as in (1) only if VP topicalization is reconstructed. Indeed, the idea that a moved predicate is reconstructed so that it can be interpreted in its base position seems to make sense from the point of view of both semantics and syntax.[11] As far as topicalized remnant NPs as in (28) (= (20) from chapter 1) are concerned, the evidence for reconstruction at LF is not so clear; however, the counter-evidence does not strike me as decisive either, so the approach based on the PBC in (15) is not called into question directly.[12]

(28)	[$_{NP}$ Ein	Buch	t$_1$]$_2$	hat	Antje	[$_{PP}$ über	die Liebe]$_1$	t$_2$	gelesen
	a	book$_{acc}$		has	Antje$_{nom}$	about	the love		read

However, the case is different with certain kinds of remnant XPs that have undergone *wh*-movement in the syntax. Here, there is semantic evidence against reconstruction of either the whole *wh*-item or a part of it. A relevant example illustrating this point is (29) (= (59-a) from chapter 1):

---

[11] See especially Heycock (1995) and Takano (1995) for syntactic arguments for predicate reconstruction based on binding-theoretic considerations.

[12] Possible counter-evidence to assuming that topicalization is always reconstructed is related to the fact that topicalization may serve a focusing function. According to standard assumptions, a focused NP cannot be interpreted in its in situ position (cf. Chomsky (1981) and Kratzer (1991), among others). If this is so, it would follow that there are contexts in which topic reconstruction may not apply at LF. The licensing of remnant topics, though, does not appear to depend on whether the topic is focused (i.e., non-reconstructable under present assumptions) or not (i.e., reconstructable).

(29) [$_{\text{NP}}$ Was für ein Buch t$_1$]$_2$ hast du [$_{\text{PP}}$ über die Liebe]$_1$ t$_2$ gelesen ?
what for a book$_{acc}$ have you about the love read

The same goes for a sentence like (30):[13]

(30) ?[$_{\text{NP}}$ Welches Buch t$_1$]$_2$ hast du [$_{\text{PP}}$ über die Liebe]$_1$ t$_2$ gelesen ?
which book$_{acc}$ have you about the love read

In both cases, PP$_1$ has undergone scrambling out of the *wh*-phrase NP$_2$, and NP$_2$ is overtly moved to SpecC, creating an unbound trace t$_1$. If this trace is to fulfill the PBC in (15), it seems that NP$_2$ or a constituent that contains t$_1$ and is dominated by NP$_2$ must be reconstructed into its base position. But there is semantic evidence that suggests that such reconstruction is not an option in (29) and (30).

Heim (1987; 1992), Stechow (1993; 1996), and Beck (1996), among others, argue that in the case of *which*-phrases, no part may undergo reconstruction. The reason for this is that *wh*-phrases are to be treated on a par with existentially quantified phrases, given a standard semantics of questions along the lines of Karttunen (1977). Quantifiers are commonly interpreted as relations between sets of individuals – the first set forms the *restriction* of the quantifier and the second the *nuclear scope* (cf. Heim (1982) for this terminology). Now, in the case of *wh*-phrases of the *which book* kind (as in (29) and (30)), *which* is interpreted as an existential quantifier, and *book* acts as the restriction of this quantifier.[14] Hence, it is crucial that the non-*wh*-part of a *which*-phrase not be reconstructed – since it forms the restrictive part of the *wh*-operator (viewed as an existential quantifier), it cannot be interpreted as such in situ, i.e., within the nuclear scope. Thus, Heim and Stechow show that a (simplified) LF representation of a question like (30) must be of the form (31), given the approach of Karttunen (1977):

(31) which [$_\alpha$ λx [x book about love]] C [$_{IP}$ [$_\beta$ λx [you read x]]]

---

[13](30) is generally felt to be somewhat worse than (29), but it is not unacceptable for most speakers. The *wh*-expressions *was für ein* in (29) and *welches* in (30) differ in that the former is indefinite and the latter definite. In line with this, I would like to contend that (29) involves a weak Specificity effect (cf., e.g., Ross (1967), Chomsky (1981, 235), Webelhuth (1992, 170), and Mahajan (1992)). Indeed, the contrast between (29) and (30) is mirrored by the following examples exhibiting extraction from NP in German in the presence of an indefinite (i-a) and a definite (i-b) determiner (cf. subsection 2.3.2 of chapter 1):

(i) a. Worüber$_1$ hast du [$_{\text{NP}}$ ein Buch t$_1$] gelesen ?
about what have you a book read
b. ??Worüber$_1$ hast du [$_{\text{NP}}$ das Buch t$_1$] gelesen ?
about what have you the book read

That said, to simplify the discussion I will neglect the difference between definiteness and indefiniteness in (29) and (30) in what follows and pretend that the two sentences have the same meaning.

[14]Note that the case is different with *wh*-items like *how many* or *how much*. Here, the *wh*-operator triggers reconstruction; see the references just cited.

Clearly, to gain something like (31) from (29)/(30), two requirements must be met. First, the trace t_1 must be interpretable at LF as "about love;" I will return to that. Second, however, it is crucial that α is external to β at LF, i.e., that no part of the fronted *wh*-item is reconstructed at LF. But then, if reconstruction is a prerequisite for remnant movement, as predicted under the PBC in (15), unbound traces in *which*-phrases should not occur at S-structure, in contrast to what is shown by the examples in (29) and (30).

To sum up, there is empirical evidence that the PBC at LF in (15) is both too strong and too weak as it stands. It is too strong because it incorrectly precludes licit remnant movement of *which*-phrases in German; and it is too weak because it does not rule out illicit *wh*-lowering in partial *wh*-movement constructions in German and in Yo-Yo movement constructions in Ewe.

4.3. The Gist of the PBC

Still, in spite of the shortcomings of the LF-based PBC just mentioned, something about the PBC in (15) seems to reflect a basic property of natural language: Elements that are semantically interpreted as variables must be bound by an operator. Given that LF is a syntactic level of representation that serves directly as the input for semantic interpretation (see, e.g., Heim (1989) and Stechow (1993), among many others), and given that semantic variable binding must involve syntactic binding (i.e., c-command plus co-indexing) at LF (see, e.g., May (1985), Barss (1986, 416)), it appears that a statement like (32) is indispensable for natural language interpretation:

(32) Variables must be bound at LF.

Note that whereas the notion of binding is to be understood syntactically in (32), in terms of c-command, the notion of 'variable' in (32) is semantic in nature, and not syntactic: It includes initial traces of NP movement (anaphoric traces in the theory of Chomsky (1981; 1982)) that are co-indexed with an operator, initial traces of A-bar movement (which belong to the class of syntactic 'variables' in standard theorizing) that are co-indexed with an operator, as well as pronouns that are co-indexed with an operator; however, it does not include intermediate traces of any kind (even if they are classified as syntactic variables) or initial traces and pronouns that are not co-indexed with an operator.

With this in mind, let us now see what consequences arise if the PBC at LF in (15) is replaced by the semantically based requirement (32). First, does (32) preclude lowering? The answer is that it does so only if the lowering operation (a) leaves behind a trace that is interpreted as a variable, and (b) is not undone by subsequent raising before LF is reached. Obviously, there is then a way out for all the illicit lowering cases discussed so far, because in these cases a scenario in which both (a) and (b) hold can in principle be avoided.[15] The kind of lowering that (32) does rule out, though, is one that applies at LF only. Thus, as observed by May (1985, 6), a statement like (32) suffices to exclude a reading of (33-a) where the quantified expression *someone* is inside the scope of the matrix verb *believe*. As shown in (33-b), such a reading could be gained by LF lowering of *someone*:[16]

(33) a. Someone$_1$ believed [$_{\text{CP}}$ [$_{\text{IP}}$ Angleton suspected Philby]]
b. *t$_1$ believed [$_{\text{CP}}$ [$_{\text{IP}}$ someone$_1$ [$_{\text{IP}}$ Angleton suspected Philby]]]

Next, let us see what consequences this has for remnant movement constructions. Clearly, if the remnant XP is reconstructed for semantic reasons, as was proposed for (most cases of) remnant topicalization above, and if this reconstruction operation does not leave behind a semantic variable (there is nothing that forces us to assume that it does), there are no problems that could arise with respect to (32). But what about the cases where I have just argued that no part of the moved remnant must be reconstructed, viz., (29) and (30) (the latter example is repeated here as (34))?

(34) ?[$_{\text{NP}}$ Welches Buch t$_1$]$_2$ hast du [$_{\text{PP}}$ über die Liebe]$_1$ t$_2$ gelesen ?
which book$_{acc}$ have you about the love read

First, note that PP$_1$ in (34) is not an operator, so t$_1$ does not have to qualify as a variable. However, recall that (34) results in an LF representation like (35):

(35) which [$_\alpha$ λx [x book about love]] C [$_{\text{IP}}$ [$_\beta$ λx [you read x]]]

Thus, it seems as though the trace t$_1$ in (34) must be interpreted as if it were its antecedent, and the antecedent may not be interpreted at all. To achieve this result, I would like to suggest adopting the copy theory of movement envisaged in Chomsky (1981) and developed in Chomsky (1993; 1995). According to this theory, traces that result from S-structure movement are full copies of their antecedents, with an articulated structure. Under the copy theory, a sentence like (34) actually looks like (36) (traces in (36) are in italics):

(36) ?[$_{\text{NP}_2}$ Welches Buch *[PP$_1$ über die Liebe]*] hast du [$_{\text{PP}_1}$ über die Liebe] *[NP$_2$ welches Buch [PP$_1$ über die Liebe]] gelesen ?*

Only the non-italicized items are pronounced at PF; the italicized parts are deleted. As far as LF is concerned, the correct result is obtained if the leftmost occurrences of NP$_2$ and PP$_1$ are interpreted as such and the other occurrences are either left without

[15] But note that a violation of (32) in an example like (17) in English can be circumvented only at the cost of a violation of other grammatical principles; recall the discussion in subsection 4.1.

[16] Note that in (33-b) t$_1$ qualifies as a semantic variable, susceptible to (32), because it is an initial trace co-indexed with a quantified expression.

semantic interpretation, or are interpreted as a variable.[17]

Clearly, the copy theory of movement also accounts for the above-mentioned cases of reconstruction with remnant topicalization – here, the fronted item is pronounced, and its trace is interpreted at LF. I take it that all instances of remnant movement can be reconciled with the requirement in (32) along these lines. However, in what follows I will generally (i.e., if nothing depends on it) abstract away from the copy theory again and continue to note traces as 't' rather than as full copies with internal structure; this is mainly for reasons of perspicuity and simplicity.[18]

To sum up this subsection, I assume that (32) is a semantically grounded constraint applying to LF representations. This constraint is compatible with the existence of remnant movement, but it does not explain why syntactic lowering is barred. Things are different with the version of the syntactic Proper Binding Condition in the following subsection.

5. Chain Binding

As we have seen, one possibility to reconcile the PBC with remnant movement is to assume that unbound traces in remnant XPs can fulfill the PBC at LF via remnant reconstruction. However, this analysis has turned out to be quite problematic. Given that genuine reconstruction cannot be the decisive factor in licensing remnant movement, it could be assumed, following Höhle (1991), that the relevant notion is what one might call a purely syntactic "pseudo-reconstruction." Indeed, an approach along these lines is suggested by Frank, Lee & Rambow (1992, 150ff). Pseudo-reconstruction can easily be understood in terms of Barss' (1984) notion of chain-binding, which is given in (37):[19]

---

[17] Two remarks. First, note that this analysis crucially depends on the option not to analyze t_1 in (34) (= the leftmost occurrence of *[PP*$_1$ *über die Liebe]* in (36)) as a semantic variable. Thus, if nothing more is said, we expect that items that must be interpreted as variables, bound by an IP-internal operator, will create illformedness if they appear within fronted *which*-phrases. (This does not necessarily rule out anaphors in these positions, given that they are not co-indexed with a quantified element.) See Engdahl (1986) and Kim (1996) for discussion of this issue and for an approach to apparent counterexamples. Second, if we adopt the copy theory, and in particular the strong form that retains copies at LF even if they are not interpreted (see below), a couple of modifications of the theory of LF legitimacy adopted so far (especially Chomsky & Lasnik's (1993) approach to trace deletion; cf. chapter 2) appear to be necessary. However, most of these modifications strike me as fairly straightforward, and I will therefore not dwell on this issue here.

[18] Much more would have to be said about a theory of reconstruction that is based on copying. For instance, it is pointed out in Kang & Müller (1994) that there are constructions in Korean and German that require reconstruction into an LF target position which can be shown to have not been used as an intermediate position by overt movement, which, if correct, poses a problem for the copy theory. I will not discuss these issues here, assuming that problems of this type can be solved in one way or another.

[19] Also compare the theory of chain accessibility sequences in Barss (1986).

(37) *Chain-Binding*:
X chain-binds Y iff X and Y are co-indexed, and
a. X c-commands Y, or
b. X c-commands a trace of Z, where Z = Y or Z contains Y.

The PBC would then have to be revised to require chain-binding (instead of binding) at S-structure, as in (38):

(38) *The Proper Chain-Binding Condition* (PCBC):
Traces must be chain-bound at S-structure.

In all the examples involving remnants discussed so far, the traces that are unbound at S-structure (and possibly also at LF, cf. (29) and (30)) are still chain-bound at this level. The reason is that although an unbound trace in a remnant XP is not itself c-commanded by its antecedent, the latter does c-command a trace of the remnant XP, and the remnant XP contains the unbound trace. Thus, the PCBC in (38) is clearly superior to its predecessors discussed in the previous three subsections (i.e., the PBC at S-structure, the GPBC, and the PBC at LF) because it systematically permits remnant movement.[20] The question is then: Does the PCBC in (38) also exclude all the illicit cases of lowering discussed above?

At first, one might think that it does: *Wh*-lowering as in the English example (2) is barred, and the same goes for topic-lowering as in the German example (3). However, whether or not the PCBC rules out *wh*-lowering in partial *wh*-movement constructions in German (as in (23-a) & (23-b)) is not quite clear. If the scope marker is co-indexed with the *wh*-phrase undergoing partial *wh*-movement and forms an S-structure chain with it (as is standardly assumed), the PCBC incorrectly predicts that these examples are possible; thus, in order for the PCBC to rule out *wh*-lowering in scope marking constructions in German, it would be necessary to assume that the scope marker *was* cannot bind either a partially moved *wh*-phrase or its trace; and I believe that there is strong evidence against such an assumption (cf. Müller (1997)). Still more problematic is the case of Yo-Yo movement in Ewe (recall the illicit derivation in (27)). In this

---

[20] Note, incidentally, that Barss (1986, 57) and Saito (1989, 187) explicitly discuss the issue of whether a condition like (38) holds for traces or not. They come to the conclusion that whereas (a requirement like) chain-binding is valid for anaphoric expressions, it is not for traces, which they assume must be properly bound at S-structure. The evidence this claim is based on comes from ungrammatical instances of remnant *wh*-movement in English, as in (i) (instances of well-formed remnant movement in a language like German are not considered by Barss and Saito):

(i) *[$_{NP}$ Which book about t_1]$_2$ don't you know [$_{CP}$ who$_1$ to read t_2]] ?

I will address this kind of sentence in chapter 5 and will argue there that remnant *wh*-movement is in fact not systematically prohibited in English and therefore must not be excluded by a principle like the PBC at S-structure (as assumed by Barss and Saito). Rather, I will show that (i) can be excluded in the same way as illicit cases of remnant scrambling and remnant left dislocation in German (recall subsection 3.3 of chapter 1).

case, all traces respect the PCBC at S-structure because lowering is undone before Spell-Out.[21] For these reasons, I conclude that the PCBC is not the right means for treating the differences between remnant movement and syntactic lowering either.[22] More generally, then, since it seems that all versions of the Proper Binding Condition invariably run into problems and cover only a subset of the relevant data, it might be preferable to try to dispense with such a condition in toto (with the exception of the semantics-based requirement (32)); ideally, its effects should be derivable from other principles. I pursue this enterprise in the following two sections, where I consider in turn what economy constraints and cyclicity constraints have to say about lowering and remnant movement.

6. Lowering and Economy

6.1. Fewest Steps & Last Resort

Is remnant movement permitted by economy constraints, and are the illicit cases of syntactic lowering ruled out by these constraints? Assuming that both movement operations in a remnant movement construction are triggered by strong features, and that the resulting LF representation cannot be arrived at with fewer feature checking operations, Last Resort and Fewest Steps permit remnant movement. Turning next to lowering, we have already seen that although these two constraints exclude certain instances of lowering, they do not systematically exclude this operation. Thus, recall that Fewest Steps and Last Resort conspire to rule out a derivation like (12), which, if it were possible, would undermine the account of complementizer-trace effects in English. Here, adjunction to an XP in the extended projection of the verb is followed by

---

[21] The same considerations apply in the case of the illicit derivation in (9) that would undermine the account of complementizer-trace effects in English; but, as noted above, this derivation is independently ruled out by the interaction of Fewest Steps and Last Resort, and therefore does not per se argue against the PCBC.

[22] Another potential argument against the PCBC is conceptual in nature and rests on the leading idea in the minimalist program that there are no S-structure constraints. To counter this objection, one might come up with a derivational version of the PCBC, as in (i), which one might dub Derivational Proper Binding Condition:

(i) *Derivational Proper Binding Condition* (DPBC):
A trace must be bound at one stage in the derivation.

Evidently, the DPBC is the counterpart to Lasnik & Saito's (1992) GPBC discussed in section 3 above (cf. (11)), with the universal quantification implicit in the GPBC (a trace must be bound at *all* stages in the derivation) replaced by existential quantification. The DPBC permits remnant movement without mentioning S-structure, but it can easily be verified that it does not fare any better than the representational PCBC in (38) with respect to the lowering data. On the contrary, it allows lowering in partial *wh*-movement constructions irrespective of whether or not the scope marker *was* may serve as a binder – in fact, it does not preclude any lowering operation that is undone at some point later in the derivation.

substitution in SpecC, which automatically gives rise to an improper movement configuration. However, Fewest Steps and Last Resort do not exclude lowering as in (2) (*wh*-lowering in English), (3) (topic-lowering in German), (23-a)/(23-b) (*wh*-lowering in scope marking constructions in German), and (27) (focus-lowering in Ewe). To illustrate this point, I will now discuss two of these examples under this perspective. First, consider lowering in partial *wh*-movement constructions in German, as in (23-a), which is repeated here in (39-a), together with its LF representation in (39-b) (= (24)):

(39) a. *Was$_1$ meinst du [$_{CP_4}$ was$_1$ C er t$_1$ gesagt hat [$_{CP_5}$ wem$_1$ er
[+wh] think you [+wh] he said has whom$_{dat}$ he
geschlafen hat]] ?
slept has

b. wem$_1$ meinst du [$_{CP_4}$ t$_1''$ C er t$_1$ gesagt hat [$_{CP_5}$ t$_1'$ er geschlafen hat]] ?

The *wh*-phrase *wem*$_1$ (‘whom$_{dat}$’) is equipped with one [+wh] feature in the numeration. Hence, all three movement operations in (39-b) conform to Last Resort: Each step ends up in a typical checking position (SpecC) for the lowest-ranked unchecked morphological feature of *wem*$_1$ (which is [+wh], assuming that the [Case] feature of the dative NP has been checked (via dative shift) prior to the first *wh*-movement operation), even though only the last step (3) results in feature checking with the matrix [+wh] C node. Similarly, Fewest Steps is not violated by this derivation: True, there are three instances of movement applying to the *wh*-phrase, and a competing derivation that yields the same LF representation without lowering in the first step only needs two movement operations applying to *wem*$_1$; but recall that the revised Fewest Steps condition argued for above does not actually count movement operations per se, but only feature checking operations. And of course, no derivation that converges and yields the same LF involves fewer instances of feature checking. Thus, illegitimate lowering as in (39) cannot be excluded by the interaction of Fewest Steps and Last Resort; and clearly, the same considerations apply in the case of illicit focus-lowering in Ewe.

Next, consider topic-lowering, as in the German example (3), which is repeated here, too:

(40) *dass Fritz t$_1$ sagte [$_{CP}$ dem Peter$_1$ habe [$_{IP}$ keiner die Claudia
that Fritz said ART Peter$_{dat}$ has$_{subj}$ no-one$_{nom}$ ART Claudia$_{acc}$
gesehen]]
seen

Given that t$_1$ is not the trace of an operator but rather the trace of a name, it does not have to be interpreted as a variable. Hence, under the copy theory, no additional movement operation has to apply to either t$_1$ or *dem Peter*$_1$ in (40)– t$_1$ is interpreted as a name, and the lowered NP is ignored for the purpose of interpretation. Assuming that syntactic topic-lowering was triggered by a [top] feature on NP$_1$, nothing is wrong with respect to Last Resort; and given that (40) is not identical at LF to a sentence that minimally differs in that NP_1 remains in situ at S-structure (cf. (41) – the crucial difference is the presence of the lowered NP in the embedded SpecC position), Fewest Steps is not violated by (40) either.

(41) dass Fritz dem Peter$_1$ sagte [$_{CP}$ es habe [$_{IP}$ keiner die
that Fritz ART Peter$_{dat}$ said EXPL has$_{subj}$ no-one$_{nom}$ ART
Claudia gesehen]]
Claudia$_{acc}$ seen

To sum up, remnant movement is not blocked by Fewest Steps and Last Resort; but illicit lowering in the syntax cannot be ruled out in general by these economy constraints, either.

6.2. The Shortest Path Condition

6.2.1. Yo-Yo Movement

Recall from subsection 3.3 of chapter 2 the transderivational Shortest Path Condition (35), which is repeated here in (42):

(42) *Shortest Path Condition*:
If two derivations D_1 and D_2 are in the same reference set and the movement paths of D_1 are shorter than the movement paths of D_2, then D_1 is to be preferred over D_2.

As shown in chapter 2, the Shortest Path Condition can be invoked to exclude Yo-Yo movement involving intermediate focus-lowering in Ewe and similar constructions in other languages (cf. Collins (1994)). A derivation D_1 as in (27), which is repeated here again as (43-a), has longer paths than a competing derivation D_2 that moves the focused phrase *Kofi*$_1$ to its target position in the matrix clause without an intermediate step of lowering into the embedded clause; D_2 is given in (43-b):

(43) a. *Kofi$_1$ e me gble na t$_1$ [$_{CP}$ t$'_1$ be wo fo Kosi]
b. Kofi$_1$ e me gble na t$_1$ [$_{CP}$ be é fo Kosi]

Thus, D_2 blocks D_1 via the Shortest Path Condition, and this yields the result that there cannot be effects of successive cyclicity in an embedded clause if movement starts in the matrix clause and also ends there.

Given this state of affairs, one might ask whether the Shortest Path Condition can generally be held responsible for the illegitimacy of syntactic lowering, while at the same time accounting for the very option of remnant movement. And indeed, at least the German cases involving partial *wh*-lowering as in (23-a) and (23-b) follow from the Shortest Path Condition, in complete analogy to (43-a). But again, upon closer inspection the dismal finding is that the Shortest Path Condition cannot be the factor underlying the prohibition against lowering. To illustrate this, I will first show that there are cases of illicit lowering that are not excluded by the Shortest Path Condition, and I will then return to a general problem raised by this constraint that has already been noted in the previous chapter.

6.2.2. Topic-Lowering

A relevant case is again topic-lowering, as in the ungrammatical sentence (44) (= (40)):

(44) *dass Fritz t_1 sagte [$_{CP}$ dem Peter$_1$ habe [$_{IP}$ keiner die Claudia
that Fritz said ART Peter$_{dat}$ has$_{subj}$ no-one$_{nom}$ ART Claudia$_{acc}$
gesehen]]
seen

I have argued above that the object trace t_1, which is a full copy of its antecedent in the embedded SpecTop position and does not (necessarily) qualify as a semantic variable, can be interpreted as *dem Peter* in situ. This means that no additional raising operation is necessary on the way to LF in (44). Under this assumption, however, there is no way to block (44) via the Shortest Path Condition: The movement paths of the derivation generating (44) are indeed longer than those of a derivation in which *dem Peter* remains in situ at S-structure, but these two derivations do not compete since they yield different LF representations.

In conclusion, the Shortest Path Condition does not rule out all cases of illicit syntactic lowering: It fails as soon as lowering is not undone by LF raising.

6.2.3. Wh-Movement at LF

I noted in subsection 3.3.2 of chapter 2 that if there is *wh*-movement of *wh*-in situ phrases at LF, the transderivational Shortest Path Condition cannot derive superiority effects (cf. (45-a) vs. (45-b)) as assumed by Chomsky (1993, 14) and Kitahara (1993, 113ff).

(45) a. I wonder [$_{CP}$ who$_1$ C [$_{IP}$ t_1 bought what$_2$]]
b. *I wonder [$_{CP}$ what$_2$ C [$_{IP}$ who$_1$ bought t_2]]

Even if we make the non-trivial assumption that the LF outputs of (45-a) and (45-b) count as “identical” for the purposes of the definition of reference set assumed here (cf. (110) from chapter 2), the movement paths of the derivations D_1 and D_2 generating the S-structure representations (45-a) and (45-b), respectively, all have the same length at LF. This undermines an explanation of superiority effects that is based on the Shortest Path Condition.

Of course, the evidence presented in the last two subsections does not speak against the Shortest Path Condition as such (i.e., it does not prove it wrong), but it shows that this transderivational constraint is poorly motivated from an empirical point of view.[23] Given this state of affairs, I think it is preferable to dispense with this constraint altogether in the form that it takes in (42). However, in chapter 5 (section 7) I will argue that a derivational residue of the transderivational Shortest Path Condition (the MLC) can fruitfully be integrated into the Last Resort condition (cf. Chomsky (1995, 297)). For the time being, though, we are left with the situation that neither a version of the Proper Binding Condition nor any of the standard economy constraints can successfully exclude all kinds of lowering and yet still permit remnant movement.

7. Strict Cyclicity

7.1. A Reformulation of the Strict Cycle Condition and the Cycle

Abstracting away from all the versions of the Proper Binding Condition and from economy constraints like Fewest Steps, Last Resort, and the Shortest Path Condition, the difference between illicit lowering and licit remnant movement is quite obvious from an intuitive point of view: Whereas both lowering and remnant movement create unbound traces, only the former operation involves downward movement; all movement operations in the remnant movement constructions encountered so far involve upward movement. Thus, the simplest way of putting it would be to state (46), which is a quote from Chomsky (1995, 253):

(46) Movement is raising, in the specific sense defined by c-command.

It seems to me that (46) accounts for all the facts tackled in this chapter in a fairly straightforward way. The question is whether (46) must be stipulated as a property of the operation Move α, or should be derived from other assumptions (as argued, e.g., by Epstein (1995)). Here, I take the latter stand; in what follows, I would like to show that (46) follows from an appropriately revised version of the Strict Cycle Condition.[24]

---

[23]Another application of the transderivational Shortest Path Condition is given in Chomsky (1995, 328); it has to do with the illegitimacy of extraction from derived NP subjects (as in passive constructions). I will address this construction in some detail in chapter 4; it will turn out that the transderivational Shortest Path Condition is not needed for an account of this construction either (see in particular footnote 9 of chapter 4).

[24]The idea that standard strict cyclicity effects and the prohibition against lowering have a common source is also present in Chomsky (1993; 1995). Thus, Chomsky's (1993) Extension Condition on Merge (that was discussed in subsection 4.1 of chapter 2 in the context of strict cyclicity effects) also blocks lowering operations unless they involve adjunction or apply at LF. Similarly, Chomsky's (1995) assumption that strong features must be checked at the root (cf. his ex. (3) on page 234) correctly rules out a number of illicit downward movement operations, in addition to standard strict cyclicity configurations. Just like the Extension Condition, I will not adopt this requirement here, for the following reasons. First, it has nothing to say about lowering (or, indeed, strict cyclicity effects) at LF, given that LF movement is never triggered by strong features. Second, the system of economy constraints adopted here (cf. section 6 of chapter 2) permits syntactic movement without actual checking of a strong feature, as long as the feature is eventually checked before Spell-Out. However, I have argued that cases of "intermediate" lowering (without feature checking) must be ruled out too, which the assumption that strong features must be checked at the root

To this end, recall first the definitions of the Strict Cycle Condition and the Cycle adopted so far (cf. (74) and (76) from chapter 2):

(47) *Strict Cycle Condition*:
No rule can target a position that is dominated by a cyclic node and does not belong to the minimal residue of the head of this cyclic node.

(48) *Cycle*:
An XP becomes a cyclic node in the derivation iff movement to the minimal residue of its head takes place.

The basic intuition behind the Strict Cycle Condition is that those parts of phrase structure that have been "affected" by (movement) operations are not internally accessible for further (movement) operations anymore. With this in mind, consider a schematic lowering structure like (49):

(49) ... [$_{UP}$ ... t_1 ... [$_{WP}$ ... [$_{YP}$ ... [$_{ZP}$ α_1 [$_{Z'}$ Z ...]] ...] ...] ...]

According to (48), ZP becomes a cyclic node at the point in the derivation where α is substituted in SpecZ (which is part of Z's minimal residue). Assuming that no other movement operation has yet targeted (the minimal residue of) YP, WP, UP (which I take to be the minimal maximal projection dominating t_1), or some higher maximal projection in the tree, the Strict Cycle Condition is not violated by lowering of α.

But suppose now that the cycle is defined somewhat differently, in such a way that not only the ultimate XP landing site matters, but also all XPs that are part of a movement path:

(50) *Cycle* (revised):
An XP becomes a cyclic node in the derivation as soon as it is part of a movement path.

Movement paths are readily understood as sets of nodes that intervene between the starting point of movement and the target position. I will adopt the following definition:

(51) *Path*:
The Path from α to β is the set of nodes γ such that (a) and (b) hold:
a. γ is reflexively dominated by the minimal XP that dominates both α and β.
b. γ dominates α or β.

---

fails to do. Finally, this requirement relies on the idea that movement and structure-building operations may systematically alternate, which I do not assume here. That said, it should be kept in mind that the attempt to give a unified account of strict cyclicity effects and the prohibition against lowering which I will be concerned with in what follows is very similar to what is done in Chomsky (1995), and could in fact best be viewed as an alternative execution of that idea, on the basis of the general approach to movement that I have developed in chapter 2.

Now it follows that ZP, YP, WP, and UP in (49) become cyclic nodes as soon as lowering of α to SpecZ applies (UP is the minimal XP that dominates both t_1 and α_1 in (49) and turns into a cyclic node because of the assumption that domination is understood reflexively in (51-a)).[25]

Do these assumptions suffice to exclude lowering in a systematic way by the Strict Cycle Condition (47)? It turns out that this is not quite clear because there is a certain vagueness in the formulation of the Strict Cycle Condition in (47) that did not matter before, but does now. What exactly does it mean for a (movement) rule to "target" a position? If targeting a position does not yet imply actually being in that position, the Strict Cycle Condition does not rule out lowering as in (49) – it is the lowering operation itself that creates the cyclic node, so strict cyclicity must not be checked *before* movement. In other words, what is needed is that the Strict Cycle Condition be formulated in such a way that it is checked *after* each movement operation and *after* cyclic nodes have been created by the movement operation. For the sake of concreteness, assume first that the Strict Cycle Condition in (47) is reformulated as in (52):

(52) *Strict Cycle Condition* (revised):
If a landing site of movement is dominated by a cyclic node, it must belong to the minimal residue of the head of this cyclic node.

The question now is where in the derivation the Strict Cycle Condition (52) holds. Obviously, it cannot hold throughout. If it did, there could, e.g., never be a well-formed

---

[25]See Pesetsky (1982), Kayne (1984, ch. 7), Stechow & Sternefeld (1988), and Sternefeld (1997) for alternative definitions of movement path, which, however, do not immediately extend to lowering structures. Collins (1994) proposes a definition that does cover lowering. It is given in (i):

(i) *Path* (Collins (1994, 56)):
Let P_1 and P_2 be two categories in a tree. Let S_1 be the set of categories dominating P_1 and let S_2 be the set of categories dominating P_2. The path between P_1 and P_2 is defined as follows:
$\text{Path}(P_1, P_2) = (S_1 \cup S_2) - (S_1 \cap S_2)$

(i) is a notational variant of the definition given in the text in (51), with one crucial difference: The first XP dominating t_1 in (49) (i.e., UP) is not part of the movement path from t_1 to α according to (i) and consequently does not become a cyclic node by this lowering operation; in contrast, as just noted, it is part of the movement path according to (51) and therefore turns into a cyclic node after lowering. Similar considerations apply in the case of raising paths; as it stands, (i) (in contrast to (51)) predicts that the XP immediately dominating the landing site is not part of the movement path; however, we will see that there is evidence that such an XP must be part of the movement path. Furthermore, I take it that movement paths as determined under Collins' (1994) approach are somewhat counter-intuitive, since they predict that in a structure like (ii), the node marked 3 is not part of the movement path from A to B (or from B to A, for that matter), in contrast to the nodes marked 1, 2, 4, and 5.

(ii) $[_3\ [_2\ [_1$ B ...] ...] $[_4$ ... $[_5$ ... A]]]

3 dominates both A and B; hence, it is a member of the intersection of the set of nodes dominating A and the set of nodes dominating B. Since it is also a member of the union of these sets, it is not part of the movement path from A to B (or from B to A).

derivation for a sentence like (53) (= (56) from chapter 2):

(53) $[_{\text{CP}_1}$ Who knows $[_{\text{CP}_2}$ what John did]] $_?$

A derivation in which embedded *wh*-movement precedes *wh*-movement in the matrix must not violate strict cyclicity. However, if the Strict Cycle Condition holds throughout the derivation, the landing site of *what* (i.e., the embedded SpecC position) will eventually be dominated by the matrix CP, which is a cyclic node after *wh*-movement of *who*, and will therefore violate the Strict Cycle Condition. For reasons like this, I would like to suggest that the Strict Cycle Condition and the cycle are to be viewed as parts of the movement operation itself, i.e., that the operation Move α is made up of the following subcomponents, which apply in the order indicated in (54) (i.e., movement $\succ$ cycle $\succ$ strict cyclicity):[26]

(54) *Move* α:
   a. Movement of a category α to a position β, leaving a copy of α behind.
   b. Determination of cyclic nodes according to (50).
   c. Checking of the Strict Cycle Condition (52) in β.

It remains to be shown that the Strict Cycle Condition, now conceived of as part of the operation Move α, does indeed rule out all illicit cases of syntactic lowering, simultaneously permits remnant movement, and still covers all the effects attributed to strict cyclicity in chapter 2.

7.2. Lowering, Remnant Movement, and Strict Cyclicity

7.2.1. Standard Effects of Strict Cyclicity

Let me begin with a standard case showing the effects of strict cyclicity. Consider again an ungrammatical sentence involving *wh*-movement of an adjunct across a *wh*-island in English, such as (55) (= (64) from chapter 2).

(55) *How$_1$ do $[_{\text{IP}}$ you wonder $[_{\text{CP}}$ which car$_2$ $[_{\text{IP}}$ to fix t$_2$ t$_1$]]] ?

In subsection 4.1.4 of the preceding chapter, I argued that a derivation like the following should be excluded by the Strict Cycle Condition:[27]

(56) a. *D-structure*:
   $[_{\text{CP}_6}$ – do $[_{\text{IP}}$ you wonder $[_{\text{CP}_5}$ – $[_{\text{IP}}$ to fix which car$_2$ how$_1$]]]] ?

[26] To be sure, what is involved here is a case of extrinsic rule ordering, which may or may not be objectionable on conceptual grounds, under certain premises.

[27] In chapter 2, I assumed that steps (b) and (c) of (56) are one instance of Form Chain; but with Form Chain gone, we can now return to genuine successive-cyclic movement in this derivation.

b. *Movement of 'how' to the embedded SpecC*:
[$_{\text{CP}_6}$ – do [$_{\text{IP}}$ you wonder [$_{\text{CP}_5}$ how$_1$ [$_{\text{IP}}$ to fix which car$_2$ t$_1$]]]] ?
c. *Movement of 'how' to the matrix SpecC*:
[$_{\text{CP}_6}$ how$_1$ do [$_{\text{IP}}$ you wonder [$_{\text{CP}_5}$ t'_1 [$_{\text{IP}}$ to fix which car$_2$ t$_1$]]]] ?
d. *Movement of 'which car' to the embedded SpecC*:
[$_{\text{CP}_6}$ how$_1$ do [$_{\text{IP}}$ you wonder [$_{\text{CP}_5}$ which car$_2$ [$_{\text{IP}}$ to fix t$_2$ t$_1$]]]] ?

The first movement step in (b) turns the embedded CP$_5$ (which reflexively dominates itself) and all functional categories that are dominated by CP$_5$ and that dominate t$_1$ into cyclic nodes, given the definition of cyclic node in (50) and the definition of movement path in (51). According to (54), the Strict Cycle Condition (52) is checked after the determination of cyclic nodes; i.e., whether or not *how*-movement in (b) is legitimate or not according to the Strict Cycle Condition is calculated after the movement operation as such. Indeed, *how* in (b) is dominated by a cyclic node (CP$_5$) – but since it is part of the minimal residue of the head of this cyclic node, strict cyclicity is not violated by this step. This is, of course, exactly what we want – if the Strict Cycle Condition were already violated by the first movement step in (b), *wh*-movement would wrongly be expected never to be legitimate. Turning next to the second movement operation in (c), it is possible for more or less the same reason: After movement, CP$_6$ is a cyclic node (as are all the other XPs in the clause that dominate t$_1$), but *how* belongs to the minimal residue of the head of CP$_6$. The case is different with the third movement operation in (d): Raising of *which car* to the embedded SpecC position is excluded by the Strict Cycle Condition because there are cyclic nodes dominating the landing site of *which car* (viz., CP$_6$ and all other XPs that intervene between CP$_6$ and CP$_5$), and *which car* does not belong to the minimal residue of any of these cyclic nodes (except for CP$_5$ itself). Hence, embedded *wh*-movement in the last step in (56) is correctly ruled out as acyclic under present assumptions.

As before (cf. subsection 4.1.7 of chapter 2), it is necessary to assume that there is an LF cycle distinct from the S-structure cycle. This can technically be implemented in various ways. I will assume that the Spell-Out operation removes the label "cyclic node" from an XP and that the characterization of cyclic nodes according to (50) starts anew in the LF component. Under this view, the matrix CP (plus, irrelevantly, the matrix IP) is turned into a cyclic node by S-structure *wh*-movement in (57), but LF raising of the *wh*-in situ NP *what*$_3$ to the embedded SpecC position nonetheless does not violate strict cyclicity – after Spell-Out, the "cyclic node" feature is stripped from the matrix CP (and the intervening XPs), and only the embedded CP is a cyclic node, created by LF raising itself.

(57) Who$_1$ t$_1$ wonders [$_{\text{CP}}$ where$_2$ we bought what$_3$ t$_2$] $_?$

To sum up, the revised approach to strict cyclicity behaves just like the one adopted in chapter 2 with respect to standard cases. In the following subsection, I show that it also covers illicit lowering.

7.2.2. Lowering and Strict Cyclicity

I have discussed various cases of illicit lowering operations in the preceding sections: First, there is syntactic lowering that is undone after Spell-Out, as in the derivation of the English example (58-a) that ends up with the LF representation in (58-b):

(58) a. *John asked t_1 [$_{CP}$ who$_1$ Mary saw Bill]
b. who$_1$ John asked t_1 [$_{CP}$ t'_1 C$_{[}$+wh] Mary saw Bill]

Notwithstanding the fact that the English example (with the derivation given) is also ruled out by whatever is responsible for C$_{[}$+wh] requiring obligatory lexicalization of either C or SpecC at S-structure in this language, it is at least worth noting that the derivation in (58-b) is excluded by the Strict Cycle Condition in (52), given the concept of cyclic node in (50) and the internal structure of the Move α operation in (54). The lowering step (1) turns both the matrix VP and the embedded CP into a cyclic node (because these XPs are members of the movement path from t_1 to t'_1); strict cyclicity is checked afterwards, and since the position occupied by *who*$_1$ after lowering belongs to the minimal residue of the embedded C node but not to that of the matrix verb, ungrammaticality results. Subsequent raising of the *wh*-phrase restores c-command of the trace t_1, but this does not help anymore because the fatal violation has already occurred.

Exactly the same considerations apply in the German example in (59-a) with its LF in (59-b), which, as we have seen, is not independently ruled out by the effects of the "Wh-Movement Parameter."

(59) a. *Was$_1$ meinst du [$_{CP_4}$ was$_1$ C er t_1 gesagt hat [$_{CP_5}$ wem$_1$ er
[+wh] think you [+wh] he said has whom$_{dat}$ he
geschlafen hat]] ?
slept has
b. wem$_1$ meinst du [$_{CP_4}$ t''_1 C er t_1 gesagt hat [$_{CP_5}$ t'_1 er geschlafen hat]] ?

Lowering in the first step (1) turns CP$_5$ and the VP that minimally embeds it into cyclic nodes,[28] and since strict cyclicity is checked after the determination of cyclic nodes, lowering as in (59-b) is precluded.

Second, we have encountered examples where syntactic lowering is already undone before Spell-Out. This was the case in the following illegitimate derivation of an English sentence involving a complementizer-trace effect:

(60) *Who$_1$ do you think [$_{CP}$ t''_1 that [$_{IP}$ t_1 [$_{VP}$ t'_1 [$_{VP}$ left early]]]]?

[28] I abstract away from the issue of CP extraposition here and assume that CP is dominated by VP at the relevant stage of the derivation. See the following chapter, where I argue that this is indeed the case.

I have argued that (60) is ruled out by the interaction of Fewest Steps and Last Resort. But again, it is worth noting that (60) also violates the Strict Cycle Condition. Lowering from SpecI to the adjacent VP-adjoined position turns both IP and VP into cyclic nodes. As before, subsequent checking of strict cyclicity yields the result that lowering is illicit – movement ends up in a VP-adjoined position, but that position does not belong to the minimal residue of the embedding I node, as would be required by the Strict Cycle Condition.

Not surprisingly, the same situation arises in the Ewe example in (61), with the derivation as indicated:

(61) *Kofi$_1$ e me gble na t_1 $[_{\text{CP}}$ t'_1 be wo fo Kosi]

Recall that (61) cannot straightforwardly be ruled out by invoking principles of economy. However, the first *wh*-movement step (1) in this derivation immediately violates the Strict Cycle Condition. The movement path from t_1 to t'_1 consists of the matrix VP and the embedded CP, so these XPs become cyclic nodes as a result of the lowering operation. And although the first *wh*-movement operation ends up in the minimal residue of the embedded CP, this target position is still dominated by a cyclic node whose minimal residue it does not belong to, viz., by the matrix VP. Hence, strict cyclicity is violated by this lowering step, as intended.

Third, I have argued that there are illicit cases of syntactic lowering where downward movement is not undone by subsequent raising at all, neither before nor after Spell-Out. One such case was the English example (58-a), accompanied by a derivation that does not involve LF raising of the *wh*-phrase. Another one was topic-lowering, as in the German example (62):

(62) *dass Fritz t_1 sagte $[_{\text{CP}}$ dem Peter$_1$ habe $[_{\text{IP}}$ keiner die Claudia
that Fritz said ART Peter$_{dat}$ has$_{subj}$ no-one$_{nom}$ ART Claudia$_{acc}$
gesehen]]
seen

Substitution of *dem Peter*$_1$ in the embedded SpecTop position is excluded by strict cyclicity: The matrix VP, the embedded CP, and the embedded TopP are members of the movement path from t_1 to *dem Peter*$_1$, but whereas the lowered topic ends up in the minimal residue of Top, it is dominated by the cyclic nodes CP and VP, in violation of the Strict Cycle Condition.

Fourth and finally, the Strict Cycle Condition precludes cases of lowering at LF, as in the derivation (63-b), associated with an S-structure like (63-a).

(63) a. Someone$_1$ believed $[_{\text{CP}}$ $[_{\text{IP}}$ Angleton suspected Philby]]
b. *t_1 believed $[_{\text{CP}}$ $[_{\text{IP}}$ someone$_1$ $[_{\text{IP}}$ Angleton suspected Philby]]]

Quantifier lowering from the matrix SpecI position to an embedded IP-adjoined position violates the requirement that variables be bound at LF (see above); however, it also violates the Strict Cycle Condition at LF because the embedded CP, the matrix VP, and all functional projections in the matrix clause that are part of the LF movement path of *someone*$_1$ are cyclic nodes that dominate the landing site and do not contain this latter position in their minimal residues.

It goes without saying that the present approach systematically prohibits lowering at LF, even in those circumstances where lowering at LF has been argued to be necessary for reconstruction phenomena. Hence, all reconstruction phenomena must be taken care of by the copy theory (or by invoking the means of λ-conversion), and this is indeed a result that I arrived at earlier, partly for different reasons (related to transderivational economy constraints and the notion of reference set).

With illicit lowering now accounted for, let me finally turn to remnant movement and try to figure out whether or not it is legitimate under present assumptions.

7.2.3. Remnant Movement and Strict Cyclicity

A typical remnant movement example involves two S-structure movement operations. (64), e.g., involves scrambling of NP$_2$ and topicalization of VP$_1$.

(64) [$_{\text{VP}}$ t$_2$ Gelesen]$_1$ hat [$_{\text{IP}}$ [$_{\text{NP}}$ das Buch]$_2$ [$_{\text{IP}}$ keiner t$_1$]]
read has the book$_{acc}$ no-one$_{nom}$

I have presupposed – more or less tacitly, and in line with most of the literature – that the order of rule application here is *NP scrambling* $\succ$ *VP topicalization*, as in (65):

(65) a. *D-structure*:
– hat [$_{\text{IP}}$ keiner [$_{\text{VP}}$ [$_{\text{NP}}$ das Buch]$_2$ gelesen]$_1$]
b. *Scrambling of NP*$_2$ *to IP*:
– hat [$_{\text{IP}}$ [$_{\text{NP}}$ das Buch]$_2$ [$_{\text{IP}}$ keiner [$_{\text{VP}}$ t$_2$ gelesen]$_1$]]
c. *Topicalization of VP*$_1$ *to SpecTop*:
[$_{\text{VP}}$ t$_2$ gelesen]$_1$ hat [$_{\text{IP}}$ [$_{\text{NP}}$ das Buch]$_2$ [$_{\text{IP}}$ keiner t$_1$]]

Although an unbound trace is created by the second movement step, both of these movement steps are raising operations, and indeed, the Strict Cycle Condition is respected here: NP scrambling turns IP into a cyclic node;[29] TopP and CP, however, are not yet cyclic nodes. NP$_2$ therefore does indeed end up in a position dominated by a cyclic node (IP), but since it also belongs to the minimal residue of this cyclic node, strict cyclicity is respected. Subsequent remnant topicalization then creates an additional cyclic node TopP, and again, although the topic is dominated by a cyclic node in its target position, it is part of the minimal residue of that cyclic node, and everything

[29] At least, this is the case if we assume that "domination" in the definition of movement path (cf. (51)) is to be understood not as "inclusion," but rather as "non-exclusion," so that an XP may dominate something even when only one segment of the XP dominates it.

is fine.[30]

In contrast, suppose now that scrambling and topicalization were to apply in reverse order, as in (66):

(66) a. *D-structure*:
– hat [$_{IP}$ keiner [$_{VP}$ [$_{NP}$ das Buch]$_2$ gelesen]$_1$]
b. *Topicalization of VP$_1$ to SpecTop*:
[$_{VP}$ [$_{NP}$ das Buch]$_2$ gelesen]$_1$ hat [$_{IP}$ keiner t$_1$]
c. *Scrambling of NP$_2$ to IP*:
[$_{VP}$ t$_2$ gelesen]$_1$ hat [$_{IP}$ [$_{NP}$ das Buch]$_2$ [$_{IP}$ keiner t$_1$]]

Here, topicalization in the first step turns TopP and all functional projections intervening between TopP and VP into cyclic nodes. Subsequent scrambling out of the fronted VP in the second step then clearly qualifies as an instance of lowering and is ruled out by the Strict Cycle Condition: The landing site of NP$_2$ is an IP-adjoined position, i.e., a position that is dominated by (and does not belong to the minimal residue of) TopP, which has turned into a cyclic node as a result of the immediately preceding movement operation.

To sum up, not only is it the case that the Strict Cycle Condition (conceived of as part of the operation Move α) systematically rules out lowering structures and at the same time allows remnant movement structures – what is more, the Strict Cycle Condition also determines a certain order of application of movement operations in remnant movement constructions. I will return to this issue in the next chapter, where I argue that the order imposed on movement operations by the Strict Cycle Condition in constructions with incomplete category fronting is independently corroborated by evidence involving Freezing and Anti-Freezing effects. For the time being, however, I will leave it at that and draw a conclusion.

8. Conclusion

The goal of this chapter was to explain the fact that unbound traces occur in remnant movement configurations, but seem to be barred otherwise. In pursuing this goal, I first considered various versions of Fiengo's (1977) Proper Binding Condition (PBC), viz., the Proper Binding Condition at S-structure, the Generalized Proper Binding Condition, the Proper Binding Condition at LF, the Derivational Proper Binding Condition, and the Proper Chain-Binding Condition. None of these conditions proved sufficient to account for the diversity of illicit lowering constructions, though. A common property of all these versions of the Proper Binding Condition is that they require c-command

[30] V/2 movement to Top is ignored in this derivation for reasons of simplicity. But it is clear that given the Strict Cycle Condition, V/2 raising must follow NP scrambling; whether it follows or precedes topicalization is not important.

of a trace either at a certain level of representation or at certain intermediate stages in the derivation. As such, I contend that all versions of the PBC are inherently representational in this sense (even the Generalized and the Derivational Proper Binding Conditions). However, it seems to me that there is no single level of representation and no single intermediate stage in the derivation where licit and illicit occurrences of unbound traces can be captured correctly. For this reason, I have rejected approaches in terms of proper binding, and, more generally, the Proper Binding Condition altogether, except for a semantically based residue – the requirement that (semantic) variables are bound at LF.

Next, I tried to derive the contrast between illicit lowering and licit remnant movement by invoking economy principles, such as the Fewest Steps condition, the Last Resort condition, and the Shortest Path Condition. The interaction of the first two conditions has turned out not to be sufficient to preclude all illicit instances of lowering (although it does not make any wrong predictions either), and the latter condition, while inherently problematic, also does not successfully rule out all cases of lowering.

From a conceptual point of view, I think that analyses in terms of proper binding and economy miss a very simple generalization that covers all the data, namely: All movement must be raising. Thus, what seems to be relevant is whether or not movement is raising or lowering: Economy principles by themselves (as assumed here) do not necessarily differentiate between upward and downward movement (as concerns, e.g., the number of feature checking operations and the length of movement paths), and similarly, versions of the Proper Binding Condition differentiate between bound and unbound traces, but not directly between upward and downward movement.

In view of this, I have argued (essentially following Chomsky (1995)) that an appropriately revised Strict Cycle Condition yields all the desired results. Given the notion of strict cyclicity in (52), and given that the Strict Cycle Condition is checked at the end of each Move α operation, after movement and the determination of cyclic nodes (cf. (54)), lowering invariably violates the Strict Cycle Condition, whereas raising does not – even if it creates an unbound trace, as in cases of remnant movement.

Chapter 4

Anti-Freezing

1. Introduction

By and large, it seems to be the case that S-structure configurations are illicit when a trace is included in a moved XP and the antecedent of the trace is outside of XP. Basically, two kinds of approaches can be distinguished that strive to account for this Freezing effect that occurs with moved items, viz., (a) a representational account (see Browning (1991)), and (b) a derivational account (see Collins (1994)). The main purpose of this chapter is to show that the derivational account is preferable to the representational account, ultimately because the former also explains the a priori unexpected non-occurrence of Freezing effects under certain, well-defined circumstances. Most notably, these “Anti-Freezing” effects systematically show up with remnant movement; and this is the second peculiar property of remnant movement noted in section 3 of chapter 1 that calls for a principled explanation.

In the course of the discussion, I will also highlight another context where Anti-Freezing effects occur, viz., constructions involving extraction from an extraposed item. I will argue that these two Anti-Freezing effects are actually two sides of the same coin: The Anti-Freezing effect with extraposition shows up because extraposition can be remnant movement. It differs from the typical cases of remnant movement discussed up to now only insofar as it goes to the right and not to the left, and hence cannot be detected as easily in SOV languages where hierarchical positions on the right periphery are not obvious.

I will begin by illustrating the Freezing effect in section 2, on the basis of evidence taken primarily from German; it will first be shown how the data can be covered under both a representational and a derivational approach to movement theory. Only after this do I return to the Anti-Freezing effect with remnant movement in German in section 3. Section 4 covers the Anti-Freezing effects arising with extraposition in German and other languages. Section 5 provides additional evidence for the proposed analysis. Finally, sections 6 and 7 highlight some consequences of the approach to (Anti-) Freezing; section 6 discusses extraction from multiply embedded clauses, and section 7 is concerned with the (non-) islandhood of the antecedent of an unbound trace in remnant movement constructions.

2. Freezing Effects

2.1. The Phenomenon

In general, extraction of an XP α from another XP β is possible only if β is in situ, as schematically depicted in (1-a), and not if it has undergone movement itself, as shown in (1-b).[1]

(1) a. ... α_1 ... $[_\beta$ ... t_1 ... $]_2$
b. *... α_1 ... $[_\beta$ ... t_1 ... $]_2$ ... t_2

Henceforth, I will refer to this as the Freezing effect.[2] Freezing effects show up regularly with all kinds of movement in German.

2.1.1. A-Movement

In (2) a Freezing effect can be observed with A-movement in passive constructions (which is optional in German); see den Besten (1981, 102ff) and Grewendorf (1989, 42ff), among others:

(2) a. Worüber$_1$ ist von keinem $[_{\text{NP}}$ ein Buch t_1 $]_2$ gelesen worden ?
about-what is by no-one a book$_{nom}$ read been
b. *Worüber$_1$ ist $[_{\text{NP}}$ ein Buch t_1 $]_2$ von keinem t_2 gelesen worden ?
about-what is a book$_{nom}$ by no-one read been

In (2-a) the nominative NP remains in situ in the VP, and *wh*-extraction of *worüber* is possible. In (2-b), on the other hand, the subject NP has undergone optional movement out of the VP into the SpecI position, and *wh*-extraction is not allowed in this configuration.

Similarly, A-movement induces a Freezing effect in English. This is shown by the following example (see, e.g., Lasnik & Saito (1992, 196) and Collins (1994, 48)):

(3)?*Who$_1$ do you think that $[_{\text{NP}}$ pictures of t_1 $]_2$ were painted t_2 ?

---

[1] Linear order is irrelevant in these abstract representations.

[2] Compare Ross's (1967/1986, 173) Frozen Structure Constraint and Wexler & Culicover's (1980, 119) Freezing Principle. Note, however, that according to Wexler & Culicover's (and also partially according to Ross's) assumptions, there are more Freezing contexts than just the one identified by (1-b); in this chapter, however, I will stick to the more narrowly defined notion of Freezing exemplified by structures such as (1-b).

The only difference to the German case is that NP raising is obligatory in passive constructions in English; so an example corresponding to (2-a) in German is ungrammatical in English, but because of Case reasons (or reasons related to the Extended Projection Principle (EPP); see Chomsky (1982; 1995)) and not because of a Freezing effect:

(4) *Who$_1$ do you think that – were painted [$_{\text{NP}}$ pictures of t$_1$]$_2$?

2.1.2. Scrambling

The same picture emerges with scrambling in German. The sentences in (5) illustrate that scrambling of an object NP (that is transparent for extraction in situ) in front of an adverbial induces a Freezing effect; i.e., it prohibits extraction.

(5) a. [$_{\text{PP}}$ Über wen]$_1$ hat der Fritz letztes Jahr [$_{\text{NP}}$ ein Buch t$_1$]$_2$
about whom has ART Fritz$_{nom}$ last year a book$_{acc}$
geschrieben ?
written

b. *[$_{\text{PP}}$ Über wen]$_1$ hat der Fritz [$_{\text{VP}}$ [NP ein Buch t$_1$]$_2$ [$_{\text{VP}}$ letztes Jahr
about whom has ART Fritz$_{nom}$ a book$_{acc}$ last year
t$_2$ geschrieben]] ?
written

An identical situation arises with scrambling in front of the subject – even if the NP is transparent for extraction in its in situ position, it turns into an island if it is scrambled. This is shown by the following pair of examples (see (53) from chapter 1, subsection 3.2):

(6) a. [$_{\text{PP}}$ Worüber]$_1$ hat keiner [$_{\text{NP}}$ ein Buch t$_1$] gelesen ?
about what has no-one$_{nom}$ a book$_{acc}$ read

b. *[$_{\text{PP}}$ Worüber]$_1$ hat [$_{\text{NP}}$ ein Buch t$_1$]$_2$ keiner t$_2$ gelesen ?
about what has a book$_{acc}$ no-one$_{nom}$ read

Similarly, *wh*-extraction of an R-pronoun from a PP in German (i.e., postposition stranding, see van Riemsdijk (1978)) becomes impossible if the PP has undergone scrambling; cf.:

(7) a. Wo$_1$ meinst du [$_{\text{CP}}$ t$'_1$ dass keiner [$_{\text{PP}}$ t$_1$ mit]$_2$ gerechnet hat] ?
what think you that no-one with counted has

b. *Wo$_1$ meinst du [$_{\text{CP}}$ t$'_1$ dass [$_{\text{PP}}$ t$_1$ mit]$_2$ keiner t$_2$ gerechnet hat] ?
what think you that with no-one counted has

2.1.3. Topicalization

Just like A-movement and scrambling, topicalization in German imposes islandhood on the topicalized item. This is shown for topicalization of a direct object NP in (8) (see Grewendorf (1988, 257), among others).

(8) a. [$_{PP}$ Über wen]$_1$ meinst du [$_{CP}$ t$'_1$ hat der Fritz [$_{NP}$ ein Buch t$_1$
about whom think you has ART Fritz$_{nom}$ a book$_{acc}$
]$_2$ geschrieben] ?
written

b. *[$_{PP}$ Über wen]$_1$ meinst du [$_{CP}$ [NP ein Buch t$_1$]$_2$ hat der
about whom think you a book$_{acc}$ has ART
Fritz t$_2$ geschrieben] ?
Fritz$_{nom}$ written

The Freezing effect that arises with topicalization of a PP in German is illustrated in (9).

(9) a. Wo$_1$ meinst du [$_{CP}$ t$'_1$ hat keiner [$_{PP}$ t$_1$ für]$_2$ gestimmt] ?
what think you has no-one for voted

b. *Wo$_1$ meinst du [$_{CP}$ [$_{PP}$ t$_1$ für]$_2$ hat keiner t$_2$ gestimmt] ?
what think you for has no-one voted

Next, consider topicalization of a VP in German; a standard example is given in (10-a). As indicated in (10-b), *wh*-extraction from this VP is – of course – possible if VP stays in situ. But again, VP topicalization induces a Freezing effect in the sense that movement out of the VP then becomes impossible; see (10-c) (these examples are identical to those in (54) in chapter 1, subsection 3.2):

(10) a. Ich denke [$_{CP}$ [$_{VP}$ das Buch gelesen]$_2$ hat keiner t$_2$]
I think the book read has no-one

b. Was$_1$ denkst du [$_{CP}$ t$'_1$ hat keiner [$_{VP}$ t$_1$ gelesen]$_2$] ?
what think you has no-one read

c. *Was$_1$ denkst du [$_{CP}$ [$_{VP}$ t$_1$ gelesen]$_2$ hat keiner t$_2$] ?
what think you read has no-one

Basically, the situation appears to be the same in English. Thus, Browning (1989, 481; 1991, 553f) notes that *wh*-movement is impossible from topicalized NPs; cf.:[3]

(11) a. Who$_1$ do you think [$_{CP}$ t$'_1$ that John would like [$_{NP}$ pictures of t$_1$]$_2$] ?
b. *Who$_1$ do you think [$_{CP}$ t$'_1$ that [$_{NP}$ pictures of t$_1$]$_2$ John would like t$_2$] ?

Similarly, Postal (1972) observes that topicalization of a PP in English results in the PP not being accessible for P-stranding anymore; cf. the contrast between PP-in situ

[3]Examples like (11-b) are judged to be only slightly deviant by Lasnik & Saito (1992, 101). (More generally, these authors deny a Freezing effect with A-bar movement.) I will follow Browning's judgement here. Note, however, that even under Lasnik and Saito's assumptions, there is a deviance with examples like (11-b) which needs to be accounted for.

(as in (12-a)) and PP topicalization (as in (12-b)):[4]

(12) a. Who$_1$ do you think [$_{CP}$ t$'_1$ he will talk [$_{PP}$ to t$_1$]$_2$] ?
b. *Who$_1$ do you think [$_{CP}$ [$_{PP}$ to t$_1$]$_2$ he will talk t$_2$] ?

2.1.4. Wh-Movement

As a final instance of Freezing effects, consider *wh*-movement. (13-a) illustrates that *wh*-extraction from a direct object *wh*-phrase in a multiple question in German is possible if the object NP occurs in situ; (13-b), in contrast, shows that *wh*-extraction from a direct object *wh*-phrase becomes ungrammatical if the *wh*-phrase occupies SpecC, i.e., if it has undergone *wh*-movement itself.

(13) a. Worüber$_1$ hast du [$_{NP}$ was für Bücher t$_1$]$_2$ gelesen ?
about what have you what for books read
b. *Worüber$_1$ hast du gesagt [$_{CP}$ [$_{NP}$ was für Bücher t$_1$]$_2$ er t$_2$ gelesen
about what have you said what for books he read
hat] ?
has

Similarly, extraction from an AP is possible only if the AP is in situ (as in (14-a)), but not if it has undergone *wh*-movement to SpecC (as in (14-b)).[5]

---

[4]P-stranding generally seems to be impossible in PP topicalization structures in the Germanic languages. However, Afrikaans might be an exception; see Plessis (1977) and Sternefeld (1994). I have no explanation to offer for this phenomenon; but see Reis (1996, 62) for critical remarks.

[5]Chomsky (1986, 26) and Chomsky & Lasnik (1993, 544) claim, on the basis of the Spanish data in (i) (from Torrego (1985)), that extraction from *wh*-moved XPs does not cross an island.

(i) a. [$_{PP}$ De que autora]$_1$ no sabes [$_{CP}$ [$_{NP}$ qué traducciones t$_1$]$_2$ han ganado [$_{IP}$ t$_2$
by what author not you-know what translations have won
premios internacionales]] ?
awards international
b. *Esta es la autora [$_{CP}$ [$_{PP}$ de la que]$_1$ [$_{IP}$ [$_{NP}$ varias traducciones t$_1$]$_2$ han ganado premios
this is the author by whom several translations have won awards
internacionales]]
international

A sentence like (i-a) at first sight appears to involve *wh*-extraction from a *wh*-moved subject NP (as opposed to *wh*-movement from a subject NP in situ in (i-b)). However, it has been argued (see Cinque (1990, 47), Sternefeld (1991, 121ff), and Müller (1995, 397-399)) that sentences like (i-a) do not really exhibit the phenomenon of extraction from a *wh*-moved NP. Rather, under this view, (i-a) involves the marked option available with *de*-phrases in the Romance languages (similarly, with *von*-phrases in German) to be base-generated as an optional argument of certain verbs (such as ‘know’ as opposed to ‘win’). This optional argument then receives an “aboutness interpretation,” in Cinque’s terms. If this is so, the generalization that *wh*-movement induces a Freezing effect remains valid.

(14) a. [$_{PP}$ Auf wen]$_1$ bist du [$_{AP}$ stolz t_1]$_2$ gewesen ?
of whom have you proud been

b. *[$_{PP}$ Auf wen]$_1$ fragst du dich [$_{AP}$ wie stolz t_1]$_2$ du sein kannst ?
of whom ask you REFL how proud you be can

2.1.5. Freezing and Barriers

Thus, given the ungrammatical sentences in (2) through (14), I take it that the correct generalization is that movement of any kind (e.g., A-movement, scrambling, topicalization, and *wh*-movement) induces a Freezing effect. In a representational fashion, this Freezing effect can be stated as follows.[6]

(15) *Freezing*:
At S-structure, a trace t may not be included in a moved XP (i.e., an XP that binds a trace) if the antecedent of t is excluded by XP.

The question arises as to how the generalization in (15) can be derived. As a first step towards this, recall the concept of barrier assumed so far (cf. (2) from chapter 2), which is repeated here:

(16) *Barrier*:
XP is a barrier for α iff there is an X^n (0≤n≤P) such that (a), (b), and (c) hold:
a. X^n includes α.
b. X^n is not a complement.
c. X^0 is distinct from Y^0, where XP is the complement of Y^0.

As shown in section 2 of chapter 2, it follows from (16) that any XP in a specifier or adjoined position is a barrier. Movement out of an XP in such a position therefore violates the Barriers Condition (cf. (1) from chapter 2), which is also repeated here, as (17):

---

[6]The clause "if the antecedent of t is excluded by XP" in (15) is necessary in the light of well-formed examples like (i-a) and (i-b) in German.

(i) a. Ich glaube [$_{CP}$ [$_{VP}$ [$_{NP}$ dem Fritz]$_1$ letztes Jahr t_1 ein Buch gegeben]$_2$ hat keiner von uns t_2]
I think ART Fritz$_{dat}$ last year a book$_{acc}$ given has no-one of us

b. Er sagte [$_{CP}$ [$_{NP}$ den Fritz]$_1$ habe keiner t_1 gesehen]$_2$ habe er t_2 gedacht
he said ART Fritz$_{acc}$ has$_{subj}$ no-one seen has$_{subj}$ he thought

In (i-a) NP scrambling takes place *within* a topicalized VP, and in (i-b) NP topicalization applies *within* a topicalized CP. In both cases a trace occurs in a moved XP (VP and CP, respectively); nevertheless, a Freezing effect fails to show up, as predicted under the formulation in (15) – in neither case is the antecedent of the trace t_1 excluded by the moved item.

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

Since no movement type may legitimately target a complement position (cf. Chomsky (1981; 1995)), it follows that any moved XP will invariably turn into a barrier, given (16). Thus, after *wh*-movement, topicalization, scrambling, or A-movement, an XP blocks extraction out of it (cf. Baltin (1984) for the basic idea). These assumptions pave the way for an account of the Freezing effects encountered so far, but closer inspection reveals that more has to be said in order to fully rule out the ungrammatical examples in (2) – (14). In the following two subsections, I will discuss two approaches to the Freezing effect (one representational and one derivational in nature), both of which incorporate the idea that XPs in derived positions turn into barriers.

2.2. A Representational Approach

The basic idea behind the representational approach (cf. Browning (1991)) is to simply look at the S-structure representation and determine whether or not a category is a barrier in its S-structure position. Thus, in a representation like (1-b), which is repeated here in (18), β is a barrier according to the definition in (16), because it can never be a complement to some head (it has to be in a specifier position or in an adjoined position).

(18) *... α_1 ... $[_\beta$ ... t_1 ... $]_2$ ... t_2

Why, then, does this barrier block movement in the representational approach? For now let us forget the theory of movement introduced in chapter 2 and assume with Browning (1991) a more traditional version of the theory of movement, as it was developed in Lasnik & Saito (1984; 1992) and Chomsky (1986). Thus, suppose that the ECP is an LF filter stating that every trace must have a feature [+γ] (i.e., [+properly governed]) at LF and that [+γ] can be assigned to a trace only under antecedent-government (otherwise, [–γ] is assigned). Suppose furthermore that the ECP and the Subjacency condition can be unified (see Cinque (1990), Sternefeld (1991), and Lasnik & Saito (1992), among others), in the sense that the Subjacency condition is also a filter barring traces which are not antecedent-governed, with two differences: First, the Subjacency filter applies at S-structure (not at LF); and second, violations of this filter at S-structure are typically weaker than violations of the ECP at LF.

The crucial assumption put forward in Lasnik & Saito (1984) and adopted in Chomsky (1986) is that the feature [$\pm\gamma$] can never be assigned in the course of the derivation, but only at some level of representation. Now suppose that (18) is an S-structure representation, with two movements (one of α and one of β) having taken place between D-structure and S-structure – this is precisely the situation in all the examples involving Freezing that I have discussed so far. Then, it follows that t_1, the trace of α, receives its γ-feature at S-structure, and since β is a barrier at S-structure, the feature will be [–γ]; i.e., the trace t_1 is then marked as not being properly governed. Since γ-marking cannot later be undone (for instance, by assigning the feature [+γ] to a [–γ]-marked trace), strong (i.e., ECP-induced) ungrammaticality arises in a structure like (18) (as long as there cannot be a β-internal intermediate trace that might antecedent-govern t_1 in (18)), and the Freezing effect is explained in a relatively straightforward way.

2.3. A Derivational Approach

A derivational account of the Freezing effect, still based on a notion of barrier like that in (16), has been suggested in Browning (1989) and is developed in some detail in Collins (1994). The central assumption put forward by Collins (1994) is that in contrast to what is postulated by Lasnik & Saito (1984; 1992), γ-marking does not necessarily take place at a level of representation, but may instead apply anywhere in the derivation. Given the derivational theory of movement introduced in chapter 2, we can dispense with the concept of γ-marking altogether and nevertheless maintain Collins' insight. According to this view (taken from Chomsky & Lasnik (1993, 546f)), a trace that is generated by a movement operation which violates a locality condition (like the Barriers Condition) receives a star *, which in turn induces strong (ECP-like) ungrammaticality if it persists at LF and weak (Subjacency-like) ungrammaticality if it can be removed by a deletion operation before LF; recall, however, that a trace can be marked * only derivationally, immediately following the movement operation that created it. Thus, if a trace is not marked * immediately after the process that generates it, it cannot induce ungrammaticality at any later stage in the derivation, even if a later step in the derivation creates a configuration where the trace is contained in a barrier. It remains to be shown how the Freezing effect can be derived under this assumption.

To rule out a sentence involving a Freezing effect on the basis of the derivational approach developed by Collins (1994), one has to ensure that every possible derivation violates some principle of grammar. In what follows, I will sketch Collins' analysis.[7] As an example I will use the sentence in (2-b), which involves a Freezing effect with A-movement in German. (2-b) is repeated here as (19).

(19) *Worüber$_1$ ist [$_{NP}$ ein Buch t_1]$_2$ von keinem t_2 gelesen worden ?
about-what is a book$_{nom}$ by no-one read been

Collins observes that there are three derivations of sentences like (19) that are a priori possible. One derivation violates the Barriers Condition, the second the Strict Cycle Condition, and the third a prohibition against "chain interleaving." I will address these three derivations in the following three subsections.

[7]Note that I will sometimes tacitly adjust irrelevant differences between the derivational theory of movement adopted by Collins (1994) and the one introduced in chapter 2.

2.3.1. Derivation No. 1: Locality

The first derivation essentially mirrors the situation in the representational approach. The D-structure representation of (19) looks roughly as in (20):[8]

(20) [$_{CP}$ – [$_{IP}$ – von keinem [$_{VP}$ [$_{NP}$ ein Buch [$_{PP}$ worüber]$_1$]$_2$ gelesen worden ist]]]

First, NP_2 undergoes raising to SpecI, creating (21).

(21) *NP Raising to SpecI*:
[$_{CP}$ – [$_{IP}$ [$_{NP}$ ein Buch [$_{PP}$ worüber]$_1$]$_2$ von keinem [$_{VP}$ t_2 gelesen worden ist]]]

The trace of NP_2, t_2, is not assigned a star because VP is not a barrier according to (16). But since NP_2 itself now occupies a specifier position (SpecI), it turns into a barrier according to (16). Consequently, subsequent *wh*-movement of PP_1 to the SpecC position, as in (22), crosses a barrier, and t_1 receives a star *:

(22) *PP Extraction to SpecC*:
[$_{CP}$ [$_{PP}$ worüber]$_1$ [$_{IP}$ [$_{NP}$ ein Buch *t_1]$_2$ von keinem [$_{VP}$ t_2 gelesen worden ist]]]

Thus, if raising precedes *wh*-movement, the Freezing effect in sentences like (19) can be reduced to a violation of the Barriers Condition.

2.3.2. Derivation No. 2: The Strict Cycle Condition

But what if raising and *wh*-movement apply to the D-structure representation in (20) (repeated here in (23)) in reverse order?

(23) [$_{CP}$ – [$_{IP}$ – von keinem [$_{VP}$ [$_{NP}$ ein Buch [$_{PP}$ worüber]$_1$]$_2$ gelesen worden ist]]]

In this case, the first step is *wh*-movement of PP_1 to the SpecC position, as shown in (24).

(24) *PP Extraction to SpecC*:
[$_{CP}$ [$_{PP}$ worüber]$_1$ [$_{IP}$ – von keinem [$_{VP}$ [$_{NP}$ ein Buch t_1]$_2$ gelesen worden ist]]]

This movement does not violate the Barriers Condition: NP_2 is still in situ in a complement position of V, and given that *lesen* (‘read’) is a verb that in principle permits extraction from NP (cf. subsection 2.3.2 of chapter 1), *wh*-movement of PP_1 from NP_2 does not cross a barrier. Subsequent raising of NP_2 to SpecI also does not violate the Barriers Condition (VP not being a barrier), so that t_2 is not marked * either:

---

[8] Here and in most of what follows, I ignore V/2 movement since it does not affect the issue under discussion.

(25) *NP Raising to SpecI*:
$[_{CP}$ $[_{PP}$ worüber $]_1$ $[_{IP}$ $[_{NP}$ ein Buch t_1 $]_2$ von keinem $[_{VP}$ t_2 gelesen worden ist]]]

Thus, in this derivation, locality requirements are met. It is here that the Strict Cycle Condition enters the picture.

In the previous chapter I proposed that the Strict Cycle Condition and the notion of the cycle be formulated as in (26) and (27), respectively:

(26) *Strict Cycle Condition*:
If a landing site of movement is dominated by a cyclic node, it must belong to the minimal residue of the head of this cyclic node.

(27) *Cycle*:
An XP becomes a cyclic node in the derivation as soon as it is part of a movement path.

Moreover, I have suggested that both the cycle and the Strict Cycle Condition should be conceived of as part of the operation Move α, as in (28):

(28) *Move* α:
  a. Movement of a category α to a position β, leaving a copy of α behind.
  b. Determination of cyclic nodes.
  c. Checking of the Strict Cycle Condition in β.

That is, according to (28), there is an extrinsic rule ordering, such that the movement operation as such is followed by a classification of XPs on the movement path as cyclic nodes, and finally, strict cyclicity is checked.

Under this conception of strict cyclicity (as well as under the conception proposed in chapter 2), NP raising to SpecI in the last step in (25) violates the Strict Cycle Condition. The reason is that the previous PP extraction operation in (24) has turned CP into a cyclic node – plus, irrelevantly, IP, VP, and NP_2. Subsequent NP raising in (25) ends up in SpecI, which is a landing site dominated by a cyclic node whose minimal residue it does not belong to (viz., CP).[9]

[9]Chomsky (1995, 328) suggests that the counter-cyclic movement operation in (25) can be excluded by the transderivational Shortest Path Condition: “Passive is the same in both; *wh*-movement is “longer” in the wrong one in an obvious sense, object being more “remote” from [Spec,CP] than subject in terms of number of XP’s crossed.” Since the Shortest Path Condition (at least in its transderivational formulation) was discarded in subsection 6.2 of chapter 3, such an account is not possible under present assumptions. But given that the relevant derivation is straightforwardly excluded by the Strict Cycle Condition which, as we have seen, receives independent motivation, a resort to the Shortest Path Condition is evidently not needed here. Note incidentally that the step in (25) can also be ruled out by two further constraints in Chomsky (1995, 234 & 365): on the one hand, by a constraint that blocks movement of any category containing a trace (which I have rejected earlier; cf. footnote 22 of chapter 1 and the examples in footnote 6 of the present chapter), and, on the other hand, by the assumption that strong features must be checked at the root (which

Summarizing so far, the derivation of a sentence like (19) that involves a Freezing effect is ruled out by the Barriers Condition if NP raising precedes *wh*-movement of PP and by the Strict Cycle Condition if *wh*-movement of PP precedes NP raising.

Let us now turn to the third derivation that is a priori possible.

2.3.3. Derivation No. 3: Chain Interleaving

Consider again the D-structure representation (23) (repeated in (29)) of the sentence in (19):

(29) $[_{CP}$ – $[_{IP}$ – von keinem $[_{VP}$ $[_{NP}$ ein Buch $[_{PP}$ worüber $]_1$ $]_2$ gelesen worden ist]]]

Now suppose that the first movement step involves the extraction of PP_1 from NP_2, as in derivation no. 2. However, this time movement does not yet go all the way to SpecC, but rather stops in a VP-adjoined position:

(30) *PP Extraction to VP*:
$[_{CP}$ – $[_{IP}$ – von keinem $[_{VP}$ $[_{PP}$ worüber $]_1$ $[_{VP}$ $[_{NP}$ ein Buch t_1 $]_2$ gelesen worden ist]]]]

Since no barrier intervenes (the object NP is still in situ, in a complement position), t_1 does not violate the Barriers Condition, as indicated. Now, in a second step, raising of NP_2 to the SpecI position takes place:

(31) *NP Raising to SpecI*:
$[_{CP}$ – $[_{IP}$ $[_{NP}$ ein Buch t_1 $]_2$ von keinem $[_{VP}$ $[_{PP}$ worüber $]_1$ $[_{VP}$ t_2 gelesen worden ist]]]]

There is no barrier that separates t_2 from its antecedent NP_2, so t_2 respects locality requirements, too. Furthermore, raising of NP_2 in (31) does not violate the Strict Cycle Condition. The reason is that VP-adjunction in (30) turns VP (and, irrelevantly, NP_2) into a cyclic node, but raising to SpecI in (31) ends up in a higher position (turning IP into a cyclic node). Finally, PP_1 undergoes *wh*-movement from the VP-adjoined position to SpecC in a third step, as in (32).

(32) *PP Movement to SpecC*:
$[_{CP}$ $[_{PP}$ worüber $]_1$ $[_{IP}$ $[_{NP}$ ein Buch t_1 $]_2$ von keinem $[_{VP}$ t'_1 $[_{VP}$ t_2 gelesen worden ist]]]]

Again, this movement respects the Barriers Condition (t'_1 in the VP-adjoined position is therefore not marked *), and it does not violate strict cyclicity either. If the derivational approach is to be successful, then this derivation must clearly also be ruled

is similar in spirit to the analysis given here; cf. footnote 24 of chapter 3).

out.[10] How, then, can this be accomplished?

Collins (1994, 47ff) observes that derivations of the kind in (29) – (32) are instantiations of *chain interleaving*, in the sense that first a part of the *wh*-chain is formed (cf. (30)), then the raising chain is formed (cf. (31)), and finally the second part of the *wh*-chain is formed (cf. (32)). Given that this kind of chain interleaving is precluded for general reasons, derivation no. 3 is also ruled out, and the Freezing effect is derived in toto.

The question arises of why chain interleaving is prohibited. Collins argues that this is due to a Fewest Steps condition of the (standard) type in (33) (= (16) from chapter 2):

(33) *Fewest Steps*:
If two derivations D_1 and D_2 are in the same reference set and D_1 involves fewer operations than D_2, then D_1 is to be preferred over D_2.

Assuming as before that reference sets are defined in terms of identity of LF output (and not via numerations, though not much depends on this in the present context), there cannot be any doubt that derivation no. 3 in (29) through (32) competes with derivations no. 1 and no. 2.[11] However – so Collins (1994) argues – derivation no. 3 is longer than the other two derivations because it involves *three* movement operations as opposed to *two* movement operations in (20) – (22) and (23) – (25). Thus, although derivation no. 3 is the only one which does not violate either the Barriers Condition or the Strict Cycle Condition, it is prohibited by Fewest Steps.

But recall now that the Fewest Steps condition in (33) necessitates the concept of Form Chain; otherwise, sentences like (34) would incorrectly be predicted to be ill formed (since a derivation without the intermediate trace would be more economical but would at the same time violate the Barriers Condition):

(34) How$_1$ do you think [$_{CP_2}$ t_1' that Bill fixed the car t_1] ?

As noted by Collins (1994), a resort to Form Chain still does not help in the case of derivation no. 3. Since movement applying to the PP_1 *worüber* is interrupted by another operation (raising of NP_2 to SpecI), there is no way that the two operations applying to *worüber* could be conceived of as one instance of Form Chain. Thus, derivation no. 3 continues to be excluded by Fewest Steps as being more costly than the other two competing derivations, even if Form Chain is adopted.

However, I suggested in section 6 of chapter 2 that the Fewest Steps condition should be revised in such a way that it does not count movement operations per se,

[10] Indeed, Browning (1991, 552) concludes from the availability of this type of derivation with typical Freezing constructions that the derivational approach is inferior to the representational approach.

[11] Recall from chapter 2 (footnote 16) that the presence or absence of intermediate traces must not be relevant to the question of identity or distinctness of the LF output of derivations.

but only those movement operations that are accompanied by feature checking. The revised Fewest Steps condition is given again in (35) (= (113) from chapter 2):

(35) *Fewest Steps*:
If two derivations D_1 and D_2 are in the same reference set and D_1 involves fewer checking operations than D_2, then D_1 is to be preferred over D_2.

I have argued in chapter 2 that if the transderivational Fewest Steps condition in (35) is accompanied by a revised notion of Last Resort, as in (36) (= (134) from chapter 2), Form Chain can be dispensed with entirely.

(36) *Last Resort*:
α is raised to a position β only if β is a typical checking position for the lowest-ranked unchecked morphological feature of α.

Thus, the question arises of whether the prohibition against chain interleaving in a derivation of the type in (29) – (32) can still be explained under the revised approach to economy constraints adopted here. If we consider the two movement operations applying to the *wh*-item *worüber*$_1$ more closely, it turns out that it can. To see this, consider the relevant derivation again:

(37) a. *D-structure*:
[$_{CP}$ – [$_{IP}$ – von keinem [$_{VP}$ [$_{NP}$ ein Buch [$_{PP}$ worüber]$_1$]$_2$ gelesen worden ist]]]
b. *Movement of PP$_1$ to VP*:
[$_{CP}$ – [$_{IP}$ – von keinem [$_{VP}$ [$_{PP}$ worüber]$_1$ [$_{VP}$ [$_{NP}$ ein Buch t$_1$]$_2$ gelesen worden ist]]]]
c. *Movement of NP$_2$ to SpecI*:
[$_{CP}$ – [$_{IP}$ [$_{NP}$ ein Buch t$_1$]$_2$ von keinem [$_{VP}$ [$_{PP}$ worüber]$_1$ [$_{VP}$ t$_2$ gelesen worden ist]]]]
d. *Movement of PP$_1$ to SpecC*:
[$_{CP}$ [$_{PP}$ worüber]$_1$ [$_{IP}$ [$_{NP}$ ein Buch t$_1$]$_2$ von keinem [$_{VP}$ t'_1 [$_{VP}$ t$_2$ gelesen worden ist]]]]

The first operation that applies to PP$_1$ is adjunction to VP, i.e., scrambling, given the assumptions above; and the second operation that applies to PP$_1$ is substitution in SpecC, i.e., *wh*-movement. Thus, what is involved here is actually a combination of scrambling to VP and *wh*-movement to SpecC applying to the same category. Therefore, it does not come as a surprise that the present system excludes (37) as involving improper movement, in more or less the same way that it derives the prohibition against *wh*-scrambling in German in examples like (38-a) (the LF representation of (38-a) is given in (38-b)):

(38) a. *Wie$_1$ hat [$_{IP}$ was$_2$ [$_{IP}$ der Fritz t$_1$ t$_2$ repariert]] ?
how has what$_{acc}$ ART Fritz fixed
b. [$_{SpecC}$ [$_{SpecC}$ wie$_1$] was$_2$] [$_{IP}$ der Fritz t$_1$ t$_2$ repariert hat]

In both cases, a PUB effect arises that can be reduced to the interaction of Fewest Steps in (35) and Last Resort in (36). Thus, in (37) there are two possible 'sub'derivations to be considered: One possibility is that a [scr] feature is absent in the numeration; then, intermediate scrambling of the *wh*-phrase violates Last Resort, but Fewest Steps is respected. Alternatively, if a [scr] feature does show up in the numeration, then intermediate scrambling fulfills Last Resort, but the derivation is ruled out by Fewest Steps because there are competing derivations that generate the same LF output with one feature checking operation fewer – viz., derivations no. 1 and no. 2. These derivations violate principles of grammar (the Barriers Condition and the Strict Cycle Condition, respectively), but assuming that they still converge, this does not preclude them from competing and, in this case, ultimately blocking less economical derivations that do not violate a grammar-internal principle.[12] Thus, the only relevant difference between illicit chain interleaving as in (37) and illicit *wh*-scrambling as in (38) is that scrambling of the *wh*-phrase is undone by *overt wh*-movement in the first case and by *covert wh*-movement in the second.

In conclusion, the revised approach to economy constraints developed in section 6 of chapter 2 and adopted here is capable of excluding illicit chain interleaving in constructions involving *wh*-extraction from NP and NP raising to SpecI, just like Collins' (1994) original approach in terms of the standard Fewest Steps condition. However, it is worth noting that the two approaches do not make identical predictions. In Collins' (1994) approach, chain interleaving is ruled out in principle. In contrast, in the approach adopted here, chain interleaving is not ruled out per se: It is barred only if it induces a PUB effect, i.e., if it gives rise to an improper movement configuration. I will return to this issue.

2.4. Conclusion

Summarizing so far, it turns out that both the representational and the derivational approaches to movement theory can account for the occurrence of Freezing effects. The question arises of whether there are conceptual and empirical arguments in favour of one approach or the other. A conceptual argument in support of the derivational approach pursued by Collins (1994) might be that it complies with the overall assumption in the minimalist program that S-structure does not exist as a genuine level of representation. Since the representational approach obviously makes reference to S-structure (it requires that γ-marking apply at this level), the derivational approach to Freezing effects might seem superior in this respect. On the other hand, the representational approach looks simpler, and hence might be considered more attractive from a conceptual point of view.

In view of this situation, I take it that empirical evidence that differentiates bet-

[12] Note, however, that the assumption that derivations that violate locality constraints can nevertheless converge is called into question in Chomsky (1995, 296-297).

ween the two approaches is called for. In the following two sections, I will consider two types of – at first sight – unexpected non-occurrences of Freezing effects, and I will argue that these Anti-Freezing effects strongly support a derivational approach to Freezing.[13]

3. Anti-Freezing Effects with Remnant Movement

3.1. The Problem

First, recall the descriptive generalization in (15) that I arrived at on the basis of the data in subsection 2.1; this generalization is repeated here in (39).

(39) *Freezing*:
At S-structure, a trace may not be included in a moved XP (i.e., an XP that binds a trace) if the antecedent of t is excluded by XP.

By and large, (39) seems to be valid. But, as noted in subsection 3.2 of chapter 1, there is one crucial case where the Freezing generalization makes wrong predictions, viz., remnant movement, which creates the following structure:

(40) ... $[_{\beta}$ ... t_1 ... $]_2$ ... α_1 ... t_2

---

[13]Collins (1994, 50) presents one empirical argument for the derivational approach. This argument is based on the following data involving PP extraposition from a subject NP in English (cf. Johnson (1985)).

(i) The case was judged.
a. Then $[_{\text{NP}}$ a lawyer t_1 $]_2$ appeared t_2 $[_{\text{PP}}$ with green eyes $]_1$
b. *Then $[_{\text{NP}}$ a lawyer t_1 $]_2$ appealed $[_{\text{PP}}$ with green eyes $]_1$

The claim is that PP extraction from a subject NP is possible with unaccusative verbs, as in (i-a), but not with unergative verbs, as in (i-b), the reason for this being that (i-a) (but not (i-b)) may have a derivation in which PP extraposition precedes raising of NP_2 from the direct object position to SpecI without locality, strict cyclicity, or economy being violated. However, data judgements are very subtle here, and various factors may interfere. Thus, the German data in (ii), which are structurally analogous, are both well formed; what is more, in this case, the sentence with the unergative verb (cf. (ii-b)) is even generally judged better than the corresponding sentence with an unaccusative verb (cf. (ii-a)).

(ii) a. Dann ist $[_{\text{NP}}$ eine Frau t_1] zur Tür herein gekommen $[_{\text{PP}}$ mit einer wunderschönen Stimme $]_1$
then is a woman_{nom} to the door into come with a beautiful voice

b. Dann hat $[_{\text{NP}}$ eine Frau t_1] angerufen $[_{\text{PP}}$ mit einer wunderschönen Stimme $]_1$
then has a woman_{nom} called with a beautiful voice

I will therefore assume that extraposition from NP is mainly constrained by extragrammatical factors and that subject NPs in general do not block rightward movement, but I will leave the question of how this can be derived open here (cf. Baltin (1983), Culicover & Rochemont (1990), and Müller (1996), among others).

Thus, typical instances of remnant topicalization as in (41) exhibit a configuration where a trace t_1 is included in a moved XP_2 at S-structure, and the antecedent of t_1 is outside of XP_2. Nevertheless, these structures are evidently well formed:

(41) a. $[_{CP}$ $[_{VP}$ t_1 Gelesen $]_2$ hat $[_{NP}$ das Buch $]_1$ keiner t_2]
read has the book no-one

b. $[_{NP}$ Ein Buch t_1 $]_2$ hat der Fritz $[_{PP}$ darüber $]_1$ t_2 geschrieben
a book has ART Fritz about that written

c. ?$[_{VP}$ $[_{PP}$ t_1 Mit] gerechnet $]_2$ hat da_1 keiner t_2
with counted has there no-one

In the same way, unbound traces created by remnant *wh*-movement are at variance with the Freezing generalization:

(42) a. $[_{NP}$ Was für Bücher t_1 $]_2$ hat er $[_{PP}$ über die Liebe $]_1$ t_2 gelesen ?
what for books has he about the love read

b. $[_{AP}$ Wie stolz t_1 $]_2$ ist der Fritz $[_{PP}$ auf seine Kinder $]_1$ gestern t_2 gewesen ?
how proud is ART Fritz of his children yesterday been

In (42-a) an NP from which scrambling of a PP has taken place is *wh*-moved; in (42-b) a remnant AP is *wh*-moved, again stranding the PP antecedent of the unbound trace in a scrambling position.

Data involving remnant scrambling are somewhat harder to find in German because, as was noted in subsection 3.3 of chapter 1 and will be discussed extensively in the following chapter, remnant scrambling is usually prohibited – at least if the antecedent of the unbound trace has also undergone scrambling, which is normally the case with remnant movement:

(43) *weil $[_{VP}$ t_1 gelesen $]_2$ $[_{NP}$ das Buch $]_1$ keiner t_2 hat
because read the book no-one has

Still, remnant scrambling is more or less acceptable in German if the antecedent of the unbound trace is a weak pronoun or clitic; cf.:[14]

---

[14]The intermediate status of (44) is probably due to the fact that a predicate has undergone scrambling. As is shown in (i), predicate scrambling is in general a marked phenomenon in German and requires specific intonation patterns (so-called I-topicalization) to be acceptable at all; see, e.g., Haider (1993) and Müller (1995).

(i) ??weil $[_{VP}$ das Buch gelesen $]_2$ keiner t_2 hat
because the book read no-one has

Here, I will abstract away from this intervening factor; what is important is that (44) is still much better than we would expect if a Freezing effect were involved. See chapter 5 for discussion.

(44) ??weil [$_{\text{VP}}$ t_1 gelesen $]_2$ es/'s$_1$ keiner t_2 hat
because read it no-one has

For the time being, it is immaterial why exactly an unbound trace in a scrambled item is illegitimate if its antecedent has undergone scrambling, but (more or less) possible if its antecedent is a clitic or weak pronoun. What is relevant, though, is that the relative wellformedness of (44) shows that unbound traces in scrambled remnants must not be precluded in general, as the result of a Freezing effect.

So there is a dilemma: Topicalization, *wh*-movement, and scrambling uniformly create Freezing effects with bound traces and Anti-Freezing effects with unbound traces. In fact, the remnant movement data closely resemble some of the ungrammatical examples involving Freezing that were discussed in subsection 2.1. Compare, e.g., the Anti-Freezing effect that arises with topicalization in the case of unbound traces in (41-a), (41-b), and (41-c) with the Freezing effect in the analogous examples in the case of bound traces in (10-c), (8-b), and (9-b), respectively. Similarly, there is a striking contrast between unbound traces in *wh*-moved XPs, which create an Anti-Freezing effect (cf. (42-a) and (42-b)), and bound traces in *wh*-moved XPs, which create a Freezing effect (cf. the analogous examples (13-b) and (14-b)); basically the same picture emerges with scrambling.

Thus, the question arises of how to account for the systematic Anti-Freezing effect with remnant movement (as in (41), (42), & (44)) under the representational and the derivational approaches to movement without sacrificing the explanation of the Freezing effect at the same time.

3.2. The Representational Approach

According to the representational approach, the Freezing effect in an S-structure representation like (1-b) (repeated here in (45-a)) is due to the fact that γ-marking of t_1 can apply only at this level and not in the course of the derivation – at S-structure, β is a barrier, and t_1 is consequently marked [–γ]. Obviously, by the same reasoning, the prediction is that a typical remnant movement structure like (40) (repeated in (45-b)) should also exhibit a Freezing effect, contrary to fact. Since, by assumption, t_1 can be γ-marked only at S-structure and not in the derivation, it will be marked [–γ] at this level, and the construction will be ruled out.

(45) a. *... α_1 ... [$_\beta$ ... t_1 ... $]_2$ ... t_2
b. ... [$_\beta$ ... t_1 ... $]_2$ ... α_1 ... t_2

Indeed, it has often been noted that remnant movement behaves exceptionally in this respect. Assuming XPs in derived positions to be barriers, the very existence of remnant movement is therefore unexplained under the approaches to movement theory in Lasnik & Saito (1984) and Chomsky (1986) where γ-marking is confined to a level of representation. Hence, if the representational approach is to be maintained in the light of the Anti-Freezing effect with remnant movement, additional assumptions must be made. One could, for instance, try to relate the non-occurrence of Freezing in (45-b) to the fact that the trace t_1 occurs unbound at S-structure here, which would violate a Proper Binding Condition at S-structure (as shown in the previous chapter). Along these lines, den Besten & Webelhuth (1987; 1990) and Rizzi (1990, 37ff) suggest that whatever accounts for t_1 occurring unbound in (45-b) might also account for the remnant category β not being a barrier – in both cases, the construction behaves as if β-movement had not taken place. A descriptively adequate generalization would be that whether a moved XP acts as a barrier for a trace t included in XP is determined by the D-structure position of XP if t is not bound at S-structure and by the S-structure position of XP if t is bound at S-structure. But, the critical remarks concerning the Proper Binding Condition made in the previous chapter nothwithstanding, this would be more or less a restatement of the facts and not a real explanation. Hence, I take it that the Anti-Freezing effect that arises with remnant movement is a problem for the representational approach to Freezing effects. Let us now turn to the derivational approach.

3.3. The Derivational Approach

3.3.1. Why Remnant Movement Does Not Create Islands

Since two instances of movement are involved in the schematic remnant movement configuration (40) (= (45-b)), at least two derivations have to be compared under the derivational approach to Freezing. In order to account for the Anti-Freezing effect, we have to show that one legitimate derivation exists. For the sake of concreteness, consider sentence (41-a), which is repeated here:

(46) $[_{CP}$ $[_{VP}$ t_1 Gelesen $]_2$ hat $[_{NP}$ das Buch $]_1$ keiner t_2 $]$
read has the book no-one

In one derivation, topicalization of VP_2 applies first. The VP-trace t_2 does not violate a locality constraint (the Barriers Condition), so it does not receive a star:

(47) *VP Topicalization*:
$[_{CP}$ $[_{VP}$ $[_{NP}$ das Buch $]_1$ gelesen $]_2$ hat keiner t_2 $]$

However, subsequent movement of NP_1 to an IP-adjoined scrambling position in the second step is ruled out for at least two reasons. First, VP_1 is not in a complement position anymore, so it has already turned into a barrier at the time when NP_1 extraction from VP applies; hence, this movement operation violates the Barriers Condition, with t_1 receiving a star * that persists at LF (t_1 being an initial trace that cannot be deleted on the way to LF) and induces strong ungrammaticality:

(48) *NP Lowering*:
$[_{CP}$ $[_{VP}$ $*t_1$ gelesen $]_2$ hat $[_{NP}$ das Buch $]_1$ keiner t_2 $]$

Second, however, recall that the movement operation in (48) is ruled out by the Strict Cycle Condition as a standard instance of syntactic lowering, given the approach developed in section 7 of chapter 3. The reason is this: The first movement operation in (47), which ends up in SpecTop, turns the TopP into a cyclic node (plus, irrelevantly, all XPs between VP and TopP). The second movement operation then is required by the Strict Cycle Condition to end up at least in the minimal residue of TopP. Clearly, however, scrambling of NP_2 to an IP-adjoined position ends up in a landing site that is dominated by the cyclic node TopP and does not belong to the latter's minimal residue. Incidentally, it turns out that since movement in (48) is lowering, the operation would be rendered illicit by the Strict Cycle Condition even if there had not been a previous VP topicalization operation which turned TopP into a cyclic node. This is so because the movement path of the lowering operation that left-adjoins NP_1 to IP consists of VP, TopP, and IP (cf. the definition of movement path in (51) in chapter 3). VP does not dominate the lowered NP_1, and IP dominates it (under the assumption that domination of two-segmental categories is defined via non-exclusion; see above), but NP_1 is in I's minimal residue; TopP, however, dominates NP_1, and NP_1 does not belong to Top's minimal residue, which induces a violation of strict cyclicity.

But suppose now that VP topicalization and NP scrambling apply in reverse order in a sentence like (46). First, NP_1 undergoes scrambling out of the VP to an IP-adjoined position. The trace t_1 left by this movement fulfills the Barriers Condition because VP is not a barrier in its in situ position:

(49) *NP Scrambling*:
$[_{CP}$ hat $[_{NP}$ das Buch $]_1$ keiner $[_{VP}$ t_1 gelesen $]_2$ $]$

Then, topicalization of the remnant category VP_2 takes place, creating an unbound trace. Again, no barrier intervenes in this movement:

(50) *VP Topicalization*:
$[_{CP}$ $[_{VP}$ t_1 gelesen $]_2$ hat $[_{NP}$ das Buch $]_1$ keiner t_2 $]$

Thus, this derivation does not violate locality. And, what is more, the derivation does not violate the Strict Cycle Condition either, in contrast to what was the case with the examples involving Freezing effects that were discussed in subsection 2.3. The reason is that VP topicalization in (50) goes to a position that is higher than the landing site of scrambling that has taken place before topicalization. By its very nature, remnant movement (i.e., movement of some XP that leads to an unbound XP-internal trace) extends the domain created by earlier movement and therefore meets the Strict Cycle Condition. More formally, NP scrambling in (49) creates a movement path consisting of IP and VP and turns these XPs into cyclic nodes. Since this movement operation ends up in the minimal residue of the highest of these cyclic nodes, strict cyclicity is respected. Given that IP is now a cyclic node, subsequent VP topicalization must end up in the minimal residue of I or of some higher head, which it does: It ends up in SpecTop, turning TopP into a cyclic node. This, in a nutshell, explains why remnant movement is an exception to the Freezing generalization (39), i.e., why an Anti-Freezing effect arises – in remnant movement configurations, extraction from an XP may take place prior to XP movement without creating a violation of strict cyclicity, which it may not in the examples exhibiting Freezing effects that were discussed in section 2 above.

Note, incidentally, that the derivation in (49) and (50) closely resembles the third derivation of Freezing effects that was addressed in subsection 2.3.3, viz., the one that involves chain interleaving. In fact, the first two steps are more or less identical: First, movement of an item α out of an XP β takes place; and then β itself undergoes movement to a higher position, thereby creating an unbound trace. The only difference is that in the case of the derivation involving chain interleaving, there is yet another operation (viz., movement of α to a position c-commanding β) which is illicit (due to the interaction of Fewest Steps and Last Resort, under the assumptions made here); and just this third illicit step is absent in the case of remnant movement.[15]

3.3.2. Why Remnant Movement Does Not Create Non-Islands

Clearly, this approach to remnant movement presupposes that the remnant category (i.e., the one bearing index 2 in the preceding examples) is not a barrier in its in situ position – otherwise, extraction of the item with index 1 would not be possible in the first place. Thus, it follows that remnant movement can never turn an island into a non-island. As shown in chapter 1, this prediction is indeed correct; the descriptive generalization arrived at there is that an unbound trace in a moved remnant XP β is possible only if a bound trace with β in situ is also possible.

Let me briefly discuss three cases from section 2 of chapter 1 that show this. First, there is a contrast between PP extraction from subject NPs and PP extraction from object NPs in German; the first kind of movement is prohibited, the second one is not (given the right choice of predicate). Compare (51-a) with (51-b):

(51) a. *dass darüber$_1$ merkwürdigerweise [$_{\text{NP}}$ ein Buch *t$_1$] den Karl beeindruckt hat
that about that strangely a book$_{nom}$ ART Karl$_{acc}$ impressed has

[15]Note that this approach to the Anti-Freezing effect with remnant movement crucially depends on the fact that we are free to choose between the illicit derivation in (47), (48), and the licit derivation in (49), (50). If, however, economy considerations would force us to abandon the second derivation and to choose the first one in certain contexts, the prediction would be that remnant movement is impossible in these contexts, i.e., that it creates a Freezing effect. In section 7 of the following chapter, I will argue that this prediction is indeed borne out – there is a context in which economy forces the choice of a derivation of a remnant movement construction that violates the Barriers Condition and the Strict Cycle Condition over a competing derivation that does not violate these conditions.

b. dass darüber$_1$ merkwürdigerweise keiner [$_{NP}$ ein Buch t$_1$]
that about that strangely no-one$_{nom}$ a book$_{acc}$
gelesen hat
read has

Given that subject NPs are barriers, and object NPs are not (necessarily), these data are accounted for – scrambling in (51-a) violates the Barriers Condition, as indicated (and *t$_1$ is not deletable on the way to LF). The contrast is mirrored in remnant movement constructions:

(52) a. *[$_{NP}$ Ein Buch *t$_1$]$_2$ hat darüber$_1$ merkwürdigerweise t$_2$ den Karl
a book has about that strangely ART Karl$_{acc}$
beeindruckt
impressed

b. [$_{NP}$ Ein Buch t$_1$]$_2$ hat darüber$_1$ merkwürdigerweise keiner t$_2$
a book has about that strangely no-one$_{nom}$
gelesen
read

The wellformedness of the remnant NP topicalization sentence (52-b) follows in the same way as that of the remnant VP topicalization example (46); a derivation in which PP scrambling from NP precedes NP topicalization does not violate either the Barriers Condition (the object NP is still in complement position at the stage of the derivation where extraction takes place) or the Strict Cycle Condition (the second movement operation lands in a higher position than the first movement operation). However, such a derivation cannot save remnant NP topicalization in (52-a). The reason is that NP$_2$ does not have to move to become a barrier; it is a barrier from the start. Hence, a derivation of (52-a) where PP scrambling precedes NP topicalization is excluded by the Barriers Condition, and a derivation where NP topicalization precedes PP scrambling violates both the Barriers Condition and the Strict Cycle Condition, as shown above.

Next, consider PPs. PPs are islands in German for extraction of full lexical NPs, but (generally) transparent for extraction of R-pronouns (see van Riemsdijk (1978)); this is shown in (53-a), in which a lexical NP is scrambled out of a PP, vs. (53-b), in which an R-pronoun is scrambled out of a PP:[16]

(53) a. *dass [$_{NP}$ dieser Sache]$_1$ keiner [$_{PP}$ mit *t$_1$] gerechnet hat
that this thing$_{g}$en no-one$_{g}$en with counted has

b. dass [$_{NP}$ da]$_1$ keiner [$_{PP}$ t$_1$ mit] gerechnet hat
that there no-one$_{g}$en with counted has

[16]Note that in order to permit extraction from PP at all, it must be ensured that the PP in examples like (53-b) occupies a complement position in VP. Cf. Frey & Tappe (1991) and Müller & Sternefeld (1994), among others, for arguments to this effect.

Suppose that the ban on extractions as in (53-a) can be made to follow from locality theory.[17] Then, the contrast between the remnant movement examples in (54-a) and (54-b) follows along the same lines:

(54) a. *[$_{\text{VP}}$ [$_{\text{PP}}$ *t$_1$ Mit] gerechnet]$_2$ hat [$_{\text{NP}}$ dieser Sache]$_1$ keiner t$_2$
with counted has this thing$_{g}$en no-one$_{nom}$

b. ?[$_{\text{VP}}$ [$_{\text{PP}}$ t$_1$ Mit] gerechnet]$_2$ hat da$_1$ keiner t$_2$
with counted has there no-one$_{nom}$

Again, if scrambling precedes topicalization in (54-b), locality requirements can be met and strict cyclicity is respected; but this ordering of movement rules does not help in the case of (54-a), where a locality constraint is violated even if scrambling precedes topicalization.

As a third and final case, consider the lexically determined contrast between coherent and incoherent infinitives, as it becomes manifest in the following data (= (41-b) and (41-a) from chapter 1):

(55) a. ?*dass [$_{\text{NP}}$ das Buch]$_1$ keiner [$_\alpha$ *t$_1$ zu lesen] abgelehnt hat
that the book$_{acc}$ no-one$_{nom}$ to read refused has

b. dass [$_{\text{NP}}$ das Buch]$_1$ keiner [$_\alpha$ t$_1$ zu lesen] versucht hat
that the book$_{acc}$ no-one$_{nom}$ to read tried has

A verb like *versuchen* (‘try’) permits scrambling out of an embedded infinitive, whereas a verb like *ablehnen* (‘refuse’) does not. This lexical variation is maintained under remnant movement, as in (56-a) vs. (56-b) (= (42-b) & (42-a) from chapter 1):

(56) a. *[$_\alpha$ *t$_1$ Zu lesen]$_3$ hat [$_{\text{NP}}$ das Buch]$_1$ keiner t$_3$ abgelehnt
to read has the book no-one refused

b. [$_\alpha$ t$_1$ Zu lesen]$_3$ hat [$_{\text{NP}}$ das Buch]$_1$ keiner t$_3$ versucht
to read has the book no-one tried

This follows in the way that by now is familiar: A derivation of (56-a) in which NP scrambling precedes α-topicalization does not help if α is already a barrier in its in situ position (cf. subsection 2.5 of chapter 1 for relevant references).

Summarizing so far, it turns out that an approach to Freezing effects based on the derivational movement theory adopted here also explains the Anti-Freezing effect with remnant movement without any additional stipulations, whereas the representational approach to Freezing does not offer an insightful account of the phenomenon and must resort to additional and otherwise unmotivated assumptions. As far as the main goals

[17] It argued in Müller (1995, ch. 2) that this follows from the Barriers Condition if additional assumptions about Case transmission and (abstract) incorporation are made. Alternatively, the confinement of P-stranding in German to extraction of R-pronouns might follow from a directionality constraint, as noted earlier in chapter 1.

of this monograph are concerned, I take it that the second of the four questions related to remnant movement that were raised in chapter 1 has received a satisfying answer: The Anti-Freezing effect with remnant movement may appear peculiar at first sight, but given the movement theory outlined in chapter 2, this is exactly what we expect.

It seems that one could stop here and move on to the next problem raised by remnant movement that was introduced in chapter 1 (viz.: Why does remnant movement show movement type asymmetries?). However, I will refrain from approaching this problem directly and will rather now turn to another kind of Anti-Freezing effect, viz., the one that systematically appears with extraposition. I do so for two reasons. First, I will argue that any instance of extraction from a finite clause in a language like German does in fact involve remnant movement of the clause, i.e., an unbound trace. And second, it turns out that the approach to extraposition to be developed in the following section provides a partial answer to the fourth question raised in section 4 of chapter 1: Why can remnant movement not create unbound intermediate traces?

4. Anti-Freezing Effects with Extraposition

4.1. The Phenomenon

Finite CP complements in SOV languages like German tend to occur post-verbally, and not pre-verbally like other arguments. This is shown in (57):

(57) a. ??dass er [$_{CP}$ dass Claudia Peter geküßt hat]$_2$ gesagt hat
that he that Claudia Peter kissed has said has

b. dass er t$_2$ gesagt hat [$_{CP}$ dass Claudia Peter geküßt hat]$_2$
that he said has that Claudia Peter kissed has

Standardly, it has been postulated that the CP V order in (57-a) is the base-generated one and that the V CP order in (57-b) is derived by an extraposition operation which right-adjoins CP to some maximal projection; cf. Bierwisch (1963), Thiersch (1978), Stechow & Sternefeld (1988), Grewendorf (1989), Büring (1995), and Büring & Hartmann (1997), among others.[18] Now, as has often been noted, extraction from a finite CP is possible only if CP appears to the right of V and never if CP is to the left of V; cf.:

(58) a. *(Ich weiß nicht) wen$_1$ er [$_{CP}$ t$'_1$ dass Claudia t$_1$ geküßt hat]$_2$ gesagt hat
I know not whom$_{acc}$ he that Claudia kissed has said has

[18]Given Last Resort, extraposition must be triggered by a strong feature, in analogy to scrambling and topicalization. For the time being, I will assume that if extraposition applies, a strong feature [extr] is indeed present. See chapter 6.

b. (Ich weiß nicht) wen$_1$ er t$_2$ gesagt hat [$_{CP}$ t'_1 dass Claudia t_1 geküßt hat]$_2$
I know not whom$_{acc}$ he said has that Claudia kissed has

There are two problems connected with extraction in (58-a) and (58-b) when one approaches V CP order in terms of extraposition. One is that (58-a), in which extraction from the purported in situ position occurs, is worse than its counterpart (57-a), in which no *wh*-extraction occurs (the latter I consider clumsy, but not ungrammatical, see chapter 6). This is unexpected if CP$_2$ in (58-a) occupies a base-position, viz., the object position of V (recall the definition of barrier in (16)); see, e.g., Webelhuth (1992, ch. 3) and Stechow & Sternefeld (1988, ch. 11). In the remainder of this chapter, I will ignore this problem. I will come back to it in section 3 of chapter 6, though, where it will become important. The other striking peculiarity of the extraction pattern in (58) is that under the extraposition analysis, we should expect (58-b) to exhibit a Freezing effect: At S-structure, the intermediate trace t'_1 is included in a moved CP, and the antecedent *wen*$_1$ is excluded by CP; this corresponds exactly to the typical Freezing configuration described in (39) above.

Thus, extraction from finite clauses in SOV languages like German exhibits an Anti-Freezing effect that is a priori unexpected (the same phenomenon occurs, inter alia, in Dutch, Hindi (cf. Mahajan (1990; 1994)), and Basque (cf. Ormazabal, Uriagereka, & Uribe-Etxebarria (1994)); this calls for an explanation. In the following two subsections, I will again compare the representational and the derivational approach to Freezing, and I will show that the Anti-Freezing effect with CP extraposition that is visible in (58-b) strongly supports the derivational theory.

4.2. The Representational Approach

Under the representational approach, *wh*-extraction in (58-b) is simply predicted to be impossible, if nothing else is said. Since, by assumption, CP$_2$ has undergone extraposition, it cannot occupy a complement position at S-structure; i.e., it cannot be a sister of some head (as to the exact landing site of extraposition, see below). Therefore, given the definition of barrier in (16), the creation of the intermediate trace t'_1 in (58-b) results in a violation of the Barriers Condition, and t'_1 should be marked [$-\gamma$] (or *). Given Chomsky and Lasnik's (1993) assumptions about uniform chains, this intermediate trace must be deleted on the way to LF, but a Subjacency-like violation should still be expected in (58-b), contrary to fact.[19]

Not least because of this well-known problem for a representational approach to

[19] Alternatively, if t'_1 were not generated in the first place, the initial t_1 would be marked [$-\gamma$] (*) (due to a Barriers Condition violation), and strong ungrammaticality would be expected, since initial traces are not LF-deletable.

movement theory and locality, it has sometimes (and especially in much recent literature) been assumed that extraposition does not exist as rightward movement at S-structure. An extraposed CP, on this view, would be base-generated in its S-structure position (see Webelhuth (1992, ch. 3), Bayer (1994, sect. 3; 1997), Haider (1994; 1995, sect. 2.2), and Wilder (1996, sect. 4) for German; Hoekstra (1983; 1984), Koster (1987, 194), and Zwart (1993, 333ff) for Dutch; and Mahajan (1994) for Hindi). Nevertheless, I will adopt the rightward movement analysis of complement extraposition in SOV languages in what follows. I will do so mainly for reasons discussed in Büring & Hartmann (1997) and Müller (1996); however, I will later provide an additional argument in favour of deriving complement extraposition by syntactic right-adjunction, rather than by base-generation.

At this point, let me repeat just two of the arguments given by Büring & Hartmann (1997) against the idea that CP can be base-generated to the right of V in languages like German and Dutch. First, an approach that treats "extraposition" phenomena in SOV languages in terms of base-generation has difficulty in ensuring that not all clauses that appear to the right of V are transparent for movement (i.e., exhibit Anti-Freezing effects). Thus, in a language like German, finite subject clauses, relative clauses, and complement clauses to nouns that are located to the right of V block extraction just as much as those in a middle-field internal position do. This is shown for subject clauses in (59), for noun complement clauses in (60), and for relative clauses in (61) (see Fanselow (1987), Grewendorf (1989), and Müller (1995), among others).

(59) a. *(Ich weiß nicht) wen$_1$ [$_{CP}$ t$'_1$ dass Claudia t$_1$ geküßt hat]$_2$ dich beeindruckte
I know not whom that Claudia kissed has you$_{acc}$ impressed

b. *(Ich weiß nicht) wen$_1$ (es) dich beeindruckte [$_{CP}$ t$'_1$ dass Claudia t$_1$ geküßt hat]$_2$
I know not whom it you$_{acc}$ impressed that Claudia kissed has

(60) a. *[$_{NP}$ Welches Buch]$_1$ hast du [$_{NP}$ ein Gerücht [$_{CP}$ t$'_1$ dass er t$_1$ gelesen hat]$_2$] gehört ?
which book have you a rumour that he read has heard

b. *[$_{NP}$ Welches Buch]$_1$ hast du [$_{NP}$ ein Gerücht] gehört [$_{CP}$ t$'_1$ dass er t$_1$ gelesen hat]$_2$?
which book have you a rumour heard that he read has

(61) a. *Wen$_1$ hast du [$_{NP}$ einen Artikel [$_{CP}$ der$_3$ t$_3$ t$_1$ beeindruckt hat]$_2$]
whom have you an article that impressed has

geschrieben ?
written

b. *Wen$_1$ hast du [$_{NP}$ einen Artikel] geschrieben [$_{CP}$ der$_3$ t$_3$ t$_1$ beeindruckt
whom have you an article written that impressed
hat]$_2$?
has

Under the base-generation approach, it is unclear why the (b)-examples in (59) – (61) are ill formed, given that the right-peripheral CPs occupy the same kind of position as, for instance, the transparent object clause in (58-b).[20]

Second, Büring & Hartmann (1997) point out that the base-generation approach rests on assumptions that are hardly compatible with a restrictive theory of projection as in Chomsky (1981; 1995). Thus, unless one is willing to adopt more intricate assumptions (like, e.g., that the particple is base-generated to the right of the finite verb and then obligatorily undergoes raising to the left), it seems that it must be postulated that in sentences like (58-b), CP is base-generated as a sister of the auxiliary *hat* ('has'), and not as a sister of the main verb *gesagt* ('said') that θ-marks it.

For reasons like these, I will continue to assume that V CP order in languages like German is derived by an extraposition operation.[21] But then, there is no denying that the Anti-Freezing effect visible in examples like (61) is not accountable for under the representational approach, at least not without introducing additional assumptions.

4.3. The Derivational Approach

4.3.1. The Problem

The typical extraction example (58-b) is repeated here again as (62).

(62) (Ich weiß nicht) wen$_1$ er t$_2$ gesagt hat [$_{CP}$ t$'_1$ dass Claudia t$_1$ geküßt hat]$_2$
I know not whom he said has that Claudia kissed has

By assumption, there are three movement operations to be considered here, viz., two instances of *wh*-movement applying to the *wh*-phrase and one instance of extraposition applying to CP. It remains to be shown that at least one of the possible combinations of the three movement operations respects all locality, cyclicity, and economy

[20] In the representational approach, it seems that one must assume that the prohibition against extraction that is at work in the (b)-examples is determined solely by lexical factors and not by structural ones; cf. Haider (1993).

[21] Another possibility that is compatible with the evidence just cited would be to assume that extraposition is the result of a PF operation; see Chomsky (1986) and especially Truckenbrodt (1994). See, however, subsection 5.2 below for an argument in support of syntactic extraposition that speaks against this option, too.

constraints. Consider first a derivation in which extraposition precedes the two *wh*-movement operations, as in (63):[22]

(63) a. *CP Extraposition to IP*:
... $[_{\text{CP}_3}$ – $[_{\text{IP}}$ $[_{\text{IP}}$ er t_2 gesagt hat I] $[_{\text{CP}_2}$ – dass Claudia wen_1 geküßt hat]]]
b. *Wh-Movement to SpecC*$_2$:
... $[_{\text{CP}_3}$ – $[_{\text{IP}}$ $[_{\text{IP}}$ er t_2 gesagt hat I] $[_{\text{CP}_2}$ wen_1 dass Claudia t_1 geküßt hat]]]
c. *Wh-Movement to SpecC*$_3$:
... $[_{\text{CP}_3}$ wen_1 $[_{\text{IP}}$ $[_{\text{IP}}$ er t_2 gesagt hat I] $[_{\text{CP}_2}$ *t_1' dass Claudia t_1 geküßt hat]]]

This derivation clearly violates the Barriers Condition in the third step. The reason is that *wh*-movement to SpecC_3 takes place from a CP in an adjoined position, i.e., from a barrier (cf. (16)). In addition, the derivation in (63) violates the Strict Cycle Condition; CP extraposition in the first step has turned the matrix IP into a cyclic node, and subsequent *wh*-movement to the embedded SpecC_2 position in the second step ends up in a position that is dominated by the matrix IP (via non-exclusion) but does not belong to the latter's minimal residue. Hence, it is evident that in order to be compatible with strict cyclicity, *wh*-movement to SpecC_2 must apply before extraposition of CP_2. These two operations are reversed in (64):

(64) a. *Wh-Movement to SpecC*$_2$:
... $[_{\text{CP}_3}$ – $[_{\text{IP}}$ er $[_{\text{CP}_2}$ wen_1 dass Claudia t_1 geküßt hat] gesagt hat I]]
b. *CP Extraposition to IP*:
... $[_{\text{CP}_3}$ – $[_{\text{IP}}$ $[_{\text{IP}}$ er t_2 gesagt hat I] $[_{\text{CP}_2}$ wen_1 dass Claudia t_1 geküßt hat]]]
c. *Wh-Movement to SpecC*$_3$:
... $[_{\text{CP}_3}$ wen_1 $[_{\text{IP}}$ $[_{\text{IP}}$ er t_2 gesagt hat I] $[_{\text{CP}_2}$ *t_1' dass Claudia t_1 geküßt hat]]]

This time, the Strict Cycle Condition is respected (no movement operation ends up in a lower position than any preceding movement operation); however, the Barriers

[22]Two remarks are due concerning (i). First, I assume throughout that I is right-peripheral in German, but not much hinges on this question in the present context. And second, for expository purposes I assume here that CP is right-adjoined to IP. If German does not have obligatory V-to-I movement at S-structure, CP might also be adjoined to the VP projected by the auxiliary *hat* ('has') (however, assuming a morphology-based theory of V-to-I movement as developed, e.g., by Roberts (1993), Rohrbacher (1994), and Vikner (1995), we might expect overt V-to-I raising to be obligatory in German, which has a relatively rich system of verbal inflection.) Right-adjunction to the VP projected by the lexical verb *gesagt* ('said') cannot be an option, though – as shown in (i), German obeys an adjacency requirement with respect to items of a verbal cluster in situ in the middle field (see, e.g., Kohrt (1974, 28) and Haider (1990, 95)):

(i) *dass er t_2 gesagt $[_{\text{CP}}$ dass Claudia Peter geküßt hat $]_2$ hat
that he said that Claudia Peter kissed has has

See Büring & Hartmann (1997) and Truckenbrodt (1994) for proposals as to why (i) is impossible.

Condition is, of course, still violated by *wh*-movement to SpecC_3 in the third step. It seems that the only derivation in which the Barriers Condition can be fulfilled is one in which all *wh*-movement applies before extraposition to IP (VP), as in (65):[23]

(65) a. *Wh-Movement to SpecC*$_2$:
... $[_{\text{CP}_3}$ – $[_{\text{IP}}$ er $[_{\text{CP}_2}$ wen$_1$ dass Claudia t$_1$ geküßt hat] gesagt hat I]]
b. *Wh-Movement to SpecC*$_3$:
... $[_{\text{CP}_3}$ wen$_1$ $[_{\text{IP}}$ er $[_{\text{CP}_2}$ t$'_1$ dass Claudia t$_1$ geküßt hat] gesagt hat I]]
c. *CP Extraposition to IP*:
... $[_{\text{CP}_3}$ wen$_1$ $[_{\text{IP}}$ $[_{\text{IP}}$ er t$_2$ gesagt hat I] $[_{\text{CP}_2}$ t$'_1$ dass Claudia t$_1$ geküßt hat]]]

This derivation does not violate the Barriers Condition; moreover, *wh*-movement does not violate the Strict Cycle Condition. However, this time, the extraposition operation in the third step violates strict cyclicity. *Wh*-movement in (65-a) has turned the matrix CP_3 into a cyclic node, and subsequent CP extraposition, analyzed here as right-adjunction to the embedded IP, ends up in a lower position (the result would be the same under the assumption that CP extraposition may target a VP-adjoined position). Hence, it seems that we end up with a dilemma – the Anti-Freezing effect with extraposition apparently cannot be accounted for under present assumptions. In the following subsection, I will suggest that this dilemma vanishes if we take the possibility that extraposition might be remnant movement into account.

4.3.2. Extraposition as Remnant Movement

Note first that there is one crucial factor that distinguishes extraposition from leftward movement types like *wh*-movement, topicalization, and scrambling, all of which induce Freezing effects – we cannot see exactly which hierarchical position an extraposed item shows up in, relative to the rest of the clause. Thus, suppose that extraposition may not only be analyzed as adjunction to VP or IP, as is commonly assumed, but also as adjunction to CP. More generally, it seems plausible to postulate that in principle right-adjunction to any kind of XP is possible. Given that right-adjunction exists at all, this is clearly the null hypothesis and therefore conceptually attractive – any restrictions on right-adjunction sites would require special justification. Thus, I would like to contend that (66) holds:[24]

---

[23]Such a derivation is in fact proposed by Mahajan (1990, 142ff) for Hindi and by Büring & Hartmann (1997) for German, as an account of the Anti-Freezing effect with clausal extraposition in these languages.

[24]For additional evidence for (66), see Müller (1996), where it is argued on the basis of empirical evidence that NP and PP also qualify as right-adjunction sites for extraposition. However, it is also argued, on the basis of theory-internal evidence, that CP is not a possible right-adjunction site in German. I believe that all pieces of indirect evidence against right-adjunction to CP in Müller (1996) can be explained away under the present approach, but, for reasons of overall coherence, I will not go into this here.

(66) *Right-Adjunction*:
Right-adjunction to any kind of XP is possible.

Under this assumption, there is yet another possible derivation for a sentence like (62), in addition to the ill-formed derivations in (63), (64), and (65), viz., the one given in (67):

(67) a. *Wh-Movement to SpecC$_2$*:
... [$_{CP_3}$ – [$_{IP}$ er [$_{CP_2}$ wen$_1$ dass Claudia t$_1$ geküßt hat] gesagt hat I]]
b. *Wh-Movement to SpecC$_3$*:
... [$_{CP_3}$ wen$_1$ [$_{IP}$ er [$_{CP_2}$ t$'_1$ dass Claudia t$_1$ geküßt hat] gesagt hat I]]
c. *CP$_2$ Extraposition to CP$_3$*:
... [$_{CP_3}$ [$_{CP_3}$ wen$_1$ [$_{IP}$ er t$_2$ gesagt hat I]] [$_{CP_2}$ t$'_1$ dass Claudia t$_1$ geküßt hat]]

As before, the Barriers Condition is not violated, because both *wh*-movement operations precede extraposition. But, what is more, extraposition to CP in (67-b) does not violate the Strict Cycle Condition either. *Wh*-movement in (67-a) has turned the matrix CP$_3$ into a cyclic node, and subsequent extraposition meets the demands of strict cyclicity because it ends up in the minimal residue of the head (C$_3$) of this cyclic node (i.e., in the CP-adjoined position). Thus, under this approach, what is responsible for the Anti-Freezing effect with extraposition of finite complement clauses is this: Extraposition may follow *wh*-movement in accordance with the Strict Cycle Condition because extraposition may end up in a position that is in the same minimal residue as the landing site of *wh*-movement – in fact, the landing site of extraposition (a CP-adjoined position, by assumption) is even slightly higher than the landing site of *wh*-movement (a SpecC position). Accordingly, the resulting structure in (67-b) involves an unbound trace, viz., t$'_1$. Hence, if this account of the Anti-Freezing effect with CP extraposition is on the right track, this means that extraction from finite clauses in SOV languages like German, Dutch, or Hindi invariably involves remnant movement.

In the following section, I will provide independent support for this analysis.

5. Additional Evidence

In this section, I will provide three pieces of evidence in support of the approach to the Anti-Freezing effect with extraposition which was developed in the previous section. The evidence concerns (a) NP and PP extraposition; (b) Freezing effects that show up with extraposition under certain, well-defined circumstances; and (c) extraposition in SVO languages like English. I will address these issues in turn.

5.1. PP and NP Extraposition

Under the approach taken here, it follows that an Anti-Freezing effect should not show up only with CP extraposition; in principle, such an effect should also occur with other kinds of XPs, such as PP and NP. This prediction does indeed seem to be borne out in German.

5.1.1. Anti-Freezing and PP Extraposition

Consider first PPs. So far I have shown that *wh*-extraction of an R-pronoun is possible if PP is in situ in a complement position (cf. (7-a), repeated here as (68-a)). If, however, leftward movement (like scrambling or topicalization) applies, whether postposition stranding is possible or not depends on whether PP contains a bound or an unbound trace of the R-pronoun. In the former case, a Freezing effect occurs (cf. (7-b), which is repeated here as (68-b)); in the latter case (i.e., with remnant movement), an Anti-Freezing configuration arises because extraction from PP may apply prior to topicalization without violating the Strict Cycle Condition; cf. (68-c) (= (41-c)):

(68) a. Wo$_1$ meinst du [$_{CP}$ t$'_1$ dass keiner [$_{PP}$ t$_1$ mit]$_2$ gerechnet hat] ?
what think you that no-one with counted has

b. *Wo$_1$ meinst du [$_{CP}$ t$'_1$ dass [$_{PP}$ t$_1$ mit]$_2$ keiner t$_2$ gerechnet hat ?]
what think you that with no-one counted has

c. ?[$_{VP}$ [$_{PP}$ t$_1$ Mit] gerechnet]$_2$ hat da$_1$ keiner t$_2$
with counted has there no-one

Now, by the same reasoning, we are led to expect that an Anti-Freezing effect may also occur with PP extraposition. And indeed, as has sometimes been observed, this is the case (cf. Geilfuß (1988, 13) and Müller (1991, 190-191), among others). To see this, note first that PP extraposition is an option in German:

(69) Es hat keiner t$_2$ gerechnet [$_{PP}$ da$_1$mit]$_2$
EXPL has no-one counted there-with

The examples in (70-a) (with *wh*-movement from PP) and (70-b) (with scrambling from PP) then show that a Freezing effect does not occur if extraction from PP and PP extraposition are combined.

(70) a. Wo$_1$ hat keiner t$_2$ gerechnet [$_{PP}$ t$_1$ mit]$_2$?
what has no-one counted with

b. dass da$_1$ keiner t$_2$ gerechnet hat [$_{PP}$ t$_1$ mit]$_2$
that there no-one counted has with

On the approach pursued here, this is so because PP can be right-adjoined to CP, so that extraction from PP may precede PP extraposition without inducing a violation of the Barriers Condition or the Strict Cycle Condition.[25,26]

5.1.2. Anti-Freezing and NP Extraposition

Basically, the situation seems to be similar in the case of NPs, although the issue is complicated by an intervening factor. Thus, recall that extraction from an object NP is possible in German (with certain governing verbs like *schreiben* ('write')) if the NP stays in situ in a complement position (as, for instance, in (8-a), repeated here as (71-a)). Leftward movement of NP creates a Freezing effect if the NP-internal trace is bound at S-structure (cf., e.g., (8-b), repeated as (71-b)) and an Anti-Freezing effect in the case of remnant NP topicalization, i.e., if the NP-internal trace is unbound; cf. (71-c) (= (41-b)).

(71) a. [$_{PP}$ Über wen]$_1$ meinst du [$_{CP}$ t_1' hat der Fritz [$_{NP}$ ein Buch t_1]$_2$
about whom think you has ART Fritz$_{nom}$ a book
geschrieben] ?
written

b. *[$_{PP}$ Über wen]$_1$ meinst du [$_{CP}$ [NP ein Buch t_1]$_2$ hat der
about whom think you a book$_{acc}$ has ART
Fritz t_2 geschrieben] ?
Fritz$_{nom}$ written

c. [$_{NP}$ Ein Buch t_1]$_2$ hat der Fritz [$_{PP}$ darüber]$_1$ t_2 geschrieben
a book has ART Fritz about that written

For reasons which are unclear to me, NP extraposition is available only with heavy NPs in German ('Heavy NP Shift') and has a marked character throughout. This is

---

[25]The construction in (70) seems to be constrained by additional factors; cf.:

(i) a. Da$_1$ hat keiner t_2 gestimmt [$_{PP}$ t_1 für/??gegen]
there has no-one voted for/against

b. Da$_1$ sind viele Leute t_2 gelaufen [$_{PP}$ t_1 auf/??unter]
there are many people walked on/under
'Many people walked on/under it (e.g., the bridge).'

All postpositions in (i) allow postposition stranding and PP extraposition in isolation, but the combination of the two processes may or may not lead to reduced acceptability. Since the syntactic context is identical for the legitimate and illegitimate cases in (i-a) and (i-b) and semantic considerations do not suggest themselves here (the antonymous postpositions being related to the main verb in roughly the same way in each case), one might speculate that phonological factors are relevant here, such that extraposition of a bare P category is optimal with a monosyllabic P, and gets worse the longer P gets. If something like this proves correct, it means that the degradedness of (i-ab) with *gegen* and *unter* as P is independent of the general approach developed here.

[26]In contrast to German, Dutch does not exhibit an identical behaviour of CP and PP with respect to extraction and extraposition: Whereas extraposed CPs are transparent for extraction in Dutch, extraposed PPs are islands; cf., e.g., van Riemsdijk (1985, 191), where this Freezing effect is noted as a problem.

shown in (72).[27]

(72) ?dass keiner je t_2 gelesen hat [$_{NP_2}$ Bücher [$_{PP_1}$ über dieses Thema] [$_{CP}$
that no-one ever read has books about this topic
die der Fritz verfaßt hat]]
that ART Fritz written has

What is relevant in the present context is that Heavy NP Shift does not noticeably decrease in acceptability if the extraposed NP contains a trace of some item in an NP-external position (which, according to my previous reasoning, is then unbound); and this is just what we are led to expect under the present set of assumptions. Consider (73):

(73) ?dass [$_{PP_1}$ über dieses Thema] keiner je t_2 gelesen hat [$_{NP_2}$ Bücher t_1 [$_{CP}$
that about this topic no-one ever read has books
die der Fritz verfaßt hat]]
that ART Fritz written has

Here, PP scrambling from NP to an IP-adjoined position can take place prior to extraposition to CP, and this combination of movement operations violates neither strict cyclicity nor locality constraints.

The same result arises with a combination of *wh*-extraction from NP and Heavy NP Shift for basically the same reasons (*wh*-movement to SpecC may precede extraposition to CP):

(74) ?[$_{PP_1}$ Über dieses Thema] hat keiner je t_2 gelesen [$_{NP_2}$ Bücher t_1 [$_{CP}$ die
about this topic has no-one ever read books that
der Fritz verfaßt hat]]
ART Fritz written has

5.2. Freezing Effects with Extraposition

The above explanation for the Anti-Freezing effect that shows up with extraposition in German has capitalized on the fact that extraposition can be taken to involve right-adjunction to CP rather than right-adjunction to VP or IP. Only under this assumption can there be a derivation of extraction preceding extraposition that does not violate the Strict Cycle Condition. This analysis makes a clear prediction. If, for some reason, we can see that the landing site of extraposition is lower than the landing site of the preceding extraction operation (e.g., if *wh*-movement out of CP is followed by CP

---

[27] Moreover, Heavy NP Shift is more acceptable if the extraposed NP is definite rather than indefinite (as in (72)). However, as noted before, definite NPs in German exhibit a Specificity effect and block extraction, quite independently of the issue of (Anti-) Freezing; cf. the references cited in subsection 2.3.2 of chapter 1. Also note that some speakers of German do not permit Heavy NP Shift at all, or only in intransitive constructions; for these speakers, the issue to be discussed here does not arise in the first place.

extraposition to VP), a Freezing effect is to be expected with extraposition, just as with leftward movement types. In what follows, I will give three arguments showing that this prediction is indeed borne out.

5.2.1. CP Extraposition to VP

As noted before (cf. footnote 22), there is an adjacency requirement in German which ensures that no item intervenes between an auxiliary and its main verb in situ in the middle field. Among other things, this precludes right-adjunction to VP in situ in cases like the following:

(75) *dass er gesagt [$_{\text{CP}}$ dass Fritz Caroline liebt] hat
that he said that Fritz Caroline loves has

However, CP extraposition to VP is an option in German if the VP undergoes topicalization; see Haider (1990), Truckenbrodt (1994), and Büring & Hartmann (1997), among others. This is shown in (76).

(76) [$_{\text{VP}_3}$ t_2 Gesagt [$_{\text{CP}_2}$ dass Fritz Caroline liebt]] hat er nicht t_3
said that Fritz Caroline loves has he not

What is more, Fanselow (1987; 1992) has noted that a VP as it appears in (76) may also undergo long-distance movement across a *wh*-island with only a weak Subjacency (i.e., Barriers Condition) violation arising; cf.:

(77) ??[$_{\text{VP}_3}$ t_2 Gesagt [$_{\text{CP}_2}$ dass Fritz Caroline liebt]] weiß ich nicht [$_{\text{CP}}$ ob er
said that Fritz Caroline loves know I not whether he
t_3 hat]
has

Now, the interesting thing to note is that this kind of long-distance topicalization (i.e., of a VP with an extraposed CP adjoined to it) becomes impossible as soon as CP_2 contains an unbound trace whose antecedent is located outside of VP_3. Thus, (78) is completely ungrammatical in German (see den Besten & Webelhuth (1990, 83) and Fanselow (1992, 6)):

(78) *[$_{\text{VP}_3}$ t_2 Gesagt [$_{\text{CP}_2}$ t'_1 dass Fritz t_1 liebt]] weiß ich nicht [$_{\text{CP}}$ wen$_1$ er t_3 hat]
said that Fritz loves know I not whom he has

The strong ungrammaticality of (78) cannot simply be due to a prohibition against remnant VP topicalization from a *wh*-island. As shown in (79), such a case is no worse than topicalization of a VP that does not contain an unbound trace:

(79) ??[$_{\text{VP}_2}$ t_1 Geküßt] weiß ich nicht [$_{\text{CP}}$ wen$_1$ sie t_2 hat]
kissed know I not whom she has

At first sight, the crucial difference between the (halfway) well-formed (79) and the impossible (78) seems to be that the remnant VP in (79) contains only an initial trace, whereas the remnant VP in (78) contains an additional intermediate trace in an extraposed CP. Indeed, recall that (78) is one of the two kinds of examples that were discussed in subsection 3.4 of chapter 1; there, I tentatively assumed that the correct empirical generalization is that unbound intermediate traces are impossible. However, given the approach to extraction from finite clauses in SOV languages outlined above, it is clear that this generalization cannot be correct – under this view, any extraction from an extraposed CP involves an unbound intermediate trace in SpecC. Thus, what is relevant in (78) cannot be the unbound trace t'_1 in an extraposed item as such, but rather the fact that the extraposition operation here differs from other, well-formed cases in that it cannot circumvent a Freezing effect.

And indeed, the illformedness of (78) follows under the approach to (Anti-) Freezing adopted here. To see this, consider the possible derivations of (78) on the basis of the (simplified) D-structure representation in (80).[28]

(80) – weiß ich nicht [$_{CP_4}$ – er [$_{VP_3}$ [$_{CP_2}$ – dass Fritz wen$_1$ liebt] gesagt] hat]

In the derivation of (78), four movement operations must take place, viz.: (a) two instances of *wh*-extraction, (b) CP_2 extraposition, and (c) VP topicalization. We can disregard from the start any derivation in which the *wh*-movement operations do not apply first. If CP_2 extraposition precedes the second *wh*-movement operation to $SpecC_4$, the Barriers Condition is violated, and if VP topicalization precedes the second *wh*-movement operation to $SpecC_4$, both the Barriers Condition and the Strict Cycle Condition are violated (VP in SpecTop qualifies as a barrier, and movement from within a category in SpecTop to a position dominated by the adjacent IP is invariably lowering).[29] But what about the following derivation?

(81) a. *Wh-Movement to SpecC*$_2$:
– weiß ich nicht [$_{CP_4}$ – er [$_{VP_3}$ [$_{CP_2}$ wen$_1$ dass Fritz t_1 liebt] gesagt] hat]
b. *Wh-Movement to SpecC*$_4$:
– weiß ich nicht [$_{CP_4}$ wen$_1$ er [$_{VP_3}$ [$_{CP_2}$ t'_1 dass Fritz t_1 liebt] gesagt] hat]
c. *Extraposition of CP*$_2$ *to VP*$_3$:
– weiß ich nicht [$_{CP_4}$ wen$_1$ er [$_{VP_3}$ [$_{VP_3}$ t_2 gesagt] [$_{CP_2}$ t'_1 dass Fritz t_1 liebt]] hat]

---

[28] I abstract away from V/2 movement of *weiß* ('know') and pretend, contrary to fact, that *weiß* is base-generated in the V/2 position; this is simply for ease of exposition. Furthermore, I ignore the question of extraposition of the least embedded clause CP_4 in the following derivation. I will return to multiple CP embedding in the following section, though.

[29] Note, however, that in order to derive *strong* ungrammaticality here, it has to be guaranteed that the unbound trace t'_1 in the fronted VP is not deleted on the way to LF (unlike bound argument traces). Incidentally, this might follow under the assumptions in Chomsky & Lasnik (1993) concerning deletion and uniform chains; if not, additional assumptions are necessary; cf. Cinque (1990) and Rizzi (1990) for discussion.

d. *Topicalization of VP*$_3$:
[$_{\text{VP}_3}$ [$_{\text{VP}_3}$ t$_2$ gesagt] [$_{\text{CP}_2}$ t$'_1$ dass Fritz t$_1$ liebt]] weiß ich nicht [$_{\text{CP}_4}$ wen$_1$ er t$_3$ hat]

Such a combination of *wh*-movement followed by CP extraposition is, of course, just what I have argued to be responsible for the systematic occurrence of Anti-Freezing effects with extraction from finite clauses in SOV languages like German. However, since CP$_2$ has to undergo topicalization together with VP in (81) in order to yield the word order in (78), it follows that extraposition of CP$_2$ must be right-adjunction to VP and cannot possibly be right-adjunction to CP$_4$.[30] Therefore, the Strict Cycle Condition is violated in (81) – *wh*-movement has turned CP$_4$ into a cyclic node (plus, irrelevantly, all XPs that intervene between CP$_4$ and the most deeply embedded VP, i.e., all XPs that are part of the two movement paths created by *wh*-movement), and subsequent CP$_2$ extraposition ends up in a position that is dominated by this cyclic node but does not belong to the minimal residue of its head.

Another derivation of (78) that must be ruled out is one in which CP extraposition and VP topicalization apply in reverse order, as in (82):

(82) a. *Wh-Movement to SpecC*$_2$:
– weiß ich nicht [$_{\text{CP}_4}$ – er [$_{\text{VP}_3}$ [$_{\text{CP}_2}$ wen$_1$ dass Fritz t$_1$ liebt] gesagt] hat]
b. *Wh-Movement to SpecC*$_4$:
– weiß ich nicht [$_{\text{CP}_4}$ wen$_1$ er [$_{\text{VP}_3}$ [$_{\text{CP}_2}$ t$'_1$ dass Fritz t$_1$ liebt] gesagt] hat]
c. *Topicalization of VP*$_3$:
[$_{\text{VP}_3}$ [$_{\text{CP}_2}$ t$'_1$ dass Fritz t$_1$ liebt] gesagt] weiß ich nicht [$_{\text{CP}_4}$ wen$_1$ er t$_3$ hat]
d. *Extraposition of CP*$_2$ *to VP*$_3$:
[$_{\text{VP}_3}$ [$_{\text{VP}_3}$ t$_2$ gesagt] [$_{\text{CP}_2}$ t$'_1$ dass Fritz t$_1$ liebt]] weiß ich nicht [$_{\text{CP}_4}$ wen$_1$ er t$_3$ hat]

Evidently, this derivation respects locality requirements like the Barriers Condition, and it also cannot be ruled out by invoking economy constraints. One might hope that at least one of the two final operations turns out to be counter-cyclic. However, topicalization of VP$_3$ in the third step is remnant movement (i.e., it creates an unbound trace of the *wh*-phrase *wen*) and does not violate the Strict Cycle Condition – it ends up in a higher position than the preceding *wh*-movement steps.[31] The question is:

---

[30] That would eventually yield (i):

(i) ?*[$_{\text{VP}_3}$ t$_2$ Gesagt] weiß ich nicht [$_{\text{CP}_4}$ wen$_1$ er t$_3$ hat] [$_{\text{CP}_2}$ t$'_1$ dass Fritz t$_1$ liebt]
said know I not whom he has that Fritz loves

which arguably is closer in status to (79) than to (78).

[31] Under present assumptions, a derivation that stops at this point, i.e., that dispenses with the last step in (82), does not violate any constraints. However, the S-structure representation corresponding to (82-c) is clearly ill formed. This is an instance of the more general problem (noted in subsection 4.1 above) that

Does subsequent right-adjunction of CP_2 to the topicalized VP violate the Strict Cycle Condition?

Intuitively, I think it should. VP topicalization turns the matrix TopP into a cyclic node, and the ensuing CP extraposition operation ends up in a subdomain of this cyclic node. The problem, however, is that CP extraposition in (82-b) does in fact respect the Strict Cycle Condition under the notion of minimal residue (or checking domain) in Chomsky (1993), which I have adopted so far (albeit in a somewhat different formulation); cf. (83) (= (75) from chapter 2):

(83) *Minimal Residue* ('Checking Domain'):
The minimal residue of a head X is the smallest set M of categories α such that:
a. X-Adj, SpecX, and XP-Adj is in M.
b. If α is in M, then α-Adj is in M.

The Strict Cycle Condition demands that if an item is dominated by a cyclic node immediately after movement, it must be part of the minimal residue of the head of this cyclic node; according to (83), if some category α is part of the minimal residue of a head β because α is the specifier of β, then another category γ that is adjoined to α is also part of the minimal residue of β. This is the situation that holds in (82-d) after VP-internal CP extraposition: VP_3 is the specifier of Top, and hence part of Top's minimal residue, as required by strict cyclicity; and CP_2 extraposition ends up in a position adjoined to VP_3, hence also in Top's minimal residue, again in accordance with strict cyclicity. To block this last step as counter-cyclic, I would therefore like to suggest that the notion of minimal residue should be understood in a stricter sense, without the option of recursion, as in (84):

(84) *Minimal Residue* (revised):
The minimal residue of a head X comprises X-Adj, SpecX, and XP-Adj, and nothing else.

Clearly, given (84), the derivation in (82) violates strict cyclicity in the last step and is therefore excluded, as desired. What is more, it turns out that this change does not affect any of the preceding reasonings, except for one, and that there is indeed evidence in favour of (84) as opposed to (83). Let me first turn to a piece of independent corroboration of (84) that was brought to my attention by Joachim Sabel (p.c.).

Consider the following two examples:

(85) a. [$_{\mathrm{CP}}$ [$_{\mathrm{VP}_2}$ Das Buch gelesen] hat [$_{\mathrm{IP}}$ Fritz$_1$ gestern t$_2$]]
the book$_{acc}$ read has Fritz$_{nom}$ yesterday
b. *[$_{\mathrm{CP}}$ [$_{\mathrm{VP}_2}$ Fritz$_1$ [$_{\mathrm{VP}_2}$ das Buch gelesen]] hat [$_{\mathrm{IP}}$ t$_1$ gestern t$_2$]]
Fritz$_{nom}$ the book$_{acc}$ read has yesterday

---

wh-extraction from a finite CP obligatorily triggers CP extraposition in German. Thus, there is a constraint that is violated if the last step in (82-d) does not apply; see chapter 6.

In (85-a) a VP consisting of a direct object and the verb has undergone topicalization, and the result is well formed. In (85-b), however, the subject NP *Fritz* also undergoes fronting, and the sentence is ungrammatical. Clearly, a derivation of (85-b) must be excluded in which first the VP is topicalized, and then the subject NP is adjoined to the topicalized VP, as an instance of scrambling, triggered by a [scr] feature on the subect NP.[32] This follows from the Strict Cycle Condition under the notion of minimal residue in (84) – VP topicalization turns TopP into a cyclic node, and subsequent scrambling of the subject NP to a VP-adjoined position ends up in a position that is dominated by TopP but is not part of the minimal residue of Top. If, however, a category adjoined to a member of the minimal residue of a head were also in the minimal residue of this head (as in Chomsky's original formulation and in (83)), this derivation of (85-b) could not be ruled out, and the sentence would be predicted to be well formed, contrary to fact.

So much for the supporting evidence. There is, however, one kind of movement that is now also precluded by the Strict Cycle Condition, viz., adjunction to SpecC, as assumed above for *wh*-movement at LF in multiple questions in languages like English and German. Consider (86-a) and its purported LF representation in (86-b):

(86) a. What$_1$ did you give t$_1$ to whom$_2$?
b. [$_{\text{CP}}$ [$_{\text{NP}}$ [$_{\text{NP}}$ what$_1$] [$_{\text{PP}}$ to whom]$_2$] C$_{[+wh]}$ you gave t$_1$ t$_2$]

According to the definition of movement path adopted here (cf. (51) from chapter 3), the movement path between PP$_2$ and its trace t$_2$ at LF includes CP – this is the minimal XP that dominates both of these items. Hence, CP turns into a cyclic node at LF (actually, it becomes a cyclic node for the second time in the derivation – recall that cyclic nodes created by S-structure movement are forgotten at LF). Under the definition of minimal residue in (83), PP$_2$ belongs to the minimal residue of C$_{[}$+wh], and everything is fine. But given the notion of minimal residue in (84), PP$_2$ is not part of the minimal residue of the C head of the cyclic node CP, and LF raising of PP$_2$ is precluded as an instance of lowering. I would like to suggest that this conse-

---

[32]There is some indirect evidence that subjects can in principle undergo scrambling in German, in contrast to what seems to be the case in, e.g., Japanese (cf. Saito (1985)). It has been argued argued that instances of weak pronoun fronting in German (and Dutch) involve movement to a designated Wackernagel position that precedes SpecI; see, e.g., Thiersch (1978), Cardinaletti & Roberts (1991), Schmidt (1992), Zwart (1993), Cardinaletti & Starke (1995), and subsection 3.1 of chapter 5 below; cf. (i-a).

(i) a. dass es$_1$ der Fritz der Maria t$_1$ gegeben hat
that it$_{acc}$ ART Fritz$_{nom}$ ART Maria$_{dat}$ given has
b. dass der Fritz$_2$ es$_1$ t$_2$ der Maria t$_1$ gegeben hat
that ART Fritz$_{nom}$ it$_{acc}$ ART Maria$_{dat}$ given has

If this is correct, then the fact that subjects may optionally precede weak object pronouns (cf. (i-b)) might be taken to indicate that subjects must be able to undergo scrambling to the XP that provides the Wackernagel position for pronouns in German.

quence is not as undesirable as it might seem at first sight. Note that the structures in (85-b) and (86-b) are virtually indistinguishable if category labels are removed, and I think this can be taken to indicate that there is indeed something wrong with the structure of (86-b). Thus, I will assume that *wh*-movement at LF in a multiple question in languages like English or German is in fact not (right-) adjunction to SpecC, but rather left-adjunction to CP (a landing site that is not available for overt scrambling in German), as in (87):[33]

(87) $[_{CP}$ $[_{PP}$ to whom $]_2$ $[_{CP}$ $[_{NP}$ what$_1$] $C_{[+wh]}$ you gave t_1 t_2]]

This approach to *wh*-movement in multiple questions raises a number of further issues because we now arrive at the general conclusion that adjunction to SpecX from within X' is barred; but as far as I can see, this does not pose unsurmountable problems.[34] I therefore finish this detour here, concluding that the more restrictive notion of minimal residue in (84) is empirically motivated by constructions involving illicit VP topicalization and does not make wrong predictions in other contexts. With this in mind, let

---

[33]Interestingly, this analysis also suggests itself from the point of view of semantic interpretation in a framework along the lines of Karttunen (1977), assuming that LF representations are maximally transparent; cf., e.g., Stechow (1993; 1996).

[34]To name just three potential problems: First, the Bulgarian data that are discussed in Rudin (1988) and that are often taken to show that right-adjunction to SpecC may occur in overt syntax (cf. chapter 2) must be reanalyzed in terms of left-adjunction to CP. Second, Irish constructions of the type in (i) (cf. McCloskey (1990, 226ff)) cannot be analyzed in terms of adjunction to SpecC:

(i) $[_{SpecC}$ Cé] $[_{PP}$ leis] $[_{C'}$ $[_{C}$ a] raibh tIJ ag caint]?
who with him C were you talk-Prog
'Who were you talking to?'

Here, the *wh*-phrase *cé* ('who') has undergone *wh*-movement to SpecC, and the PP *leis* ('with him') can optionally be fronted to a position between the *wh*-phrase and the complementizer *a*. It seems that a straightforward analysis of this construction could involve right-adjunction of the PP to the *wh*-phrase (that the *wh*-phrase is the target of movement is also suggested by the observation that PP fronting in this context depends on phonological properties of the *wh*-phrase; cf. McCloskey (1990, 227)); but this option does not exist under present assumptions. Hence, I conclude that PP fronting in (i) must end up in another kind of position. A possibility that comes to mind is left-adjunction to C, all the more so since PP fronting appears to be reanalyzable as P fronting in all the examples discussed by McCloskey.

Third, cases of optional NP-internal PP fronting in German, as in (ii) (cf. Tappe (1989), Bayer (1990; 1996), Webelhuth (1992), and Müller (1995), among others) cannot be analyzed as involving PP adjunction to SpecN (or SpecD):

(ii) a. $[_{NP}$ $[_{PP}$ von Peter $]_1$ das Auto t_1]
of Peter the car
b. $[_{NP}$ $[_{PP}$ in den Abgrund $]_1$ der Blick t_1]
into the abyss the view

Assuming PP to be left-adjoined to NP (DP) in this construction (as suggested by, e.g., Bayer (1990; 1996)) would be compatible with the theory of strict cyclicity adopted here; see, however, Müller (1995, 114-117) for arguments against such a view. I will have to leave a careful analysis of the construction in (ii) open.

me return to the original topic of this subsection, which is to show that sentences like (88) (= (78)), which show a Freezing effect with extraposition, can be ruled out by a derivational approach to movement theory that incorporates the Barriers Condition and the Strict Cycle Condition.

(88) *[$_{\mathrm{VP}_3}$ t$_2$ Gesagt [$_{\mathrm{CP}_2}$ t$'_1$ dass Fritz t$_1$ liebt]] weiß ich nicht [$_{\mathrm{CP}}$ wen$_1$ er t$_3$ hat]
said that Fritz loves know I not whom he has

I have considered in detail two derivations of (88) that are a priori conceivable – in one, the order of rule application is *wh-movement* $\succ$ *extraposition* $\succ$ *topicalization*; in the other, the order is *wh-movement* $\succ$ *topicalization* $\succ$ *extraposition*. Both these derivations are now excluded by the Strict Cycle Condition, given a non-recursive version of the notion of minimal residue. In addition, any derivation in which the second *wh*-movement operation that extracts the *wh*-phrase from CP$_2$ does not apply before CP extraposition and VP topicalization violates the Barriers Condition, and any derivation that involves an additional movement operation applying to the *wh*-phrase (thereby creating illicit chain interleaving) is ruled out by the interaction of Fewest Steps and Last Resort as an instance of improper movement. Thus, none of these derivations can lead to a well-formed sentence. In conclusion, strong ungrammaticality arises in examples like (88) in German because there is an additional (strong) Freezing effect that overrides the (weak) *wh*-island effect with remnant VP topicalization. More generally, the picture emerges that CP extraposition induces a Freezing effect (just like standard cases of leftward movement) if the landing site of the extraposed CP is *lower* than that of *wh*-movement, whereas CP extraposition induces an Anti-Freezing effect (just like typical instances of remnant movement) if the landing site of the extraposed CP is *higher* than that of *wh*-movement. These extraposition asymmetries correspond exactly to what we would expect, given the interaction of the Barriers Condition, the Strict Cycle Condition, and economy constraints (Fewest Steps & Last Resort).

5.2.2. PP Extraposition to VP

Basically the same point can be made with examples involving a combination of NP scrambling, PP extraposition, and VP topicalization. First, recall that an Anti-Freezing effect arises in the case of postposition stranding (i.e., extraction of an R-pronoun from a PP) preceding remnant VP topicalization in German, as in (41-c), which is repeated here once more as (89).

(89) ?[$_{\mathrm{VP}}$ [$_{\mathrm{PP}}$ t$_1$ Mit] gerechnet]$_2$ hat da$_1$ keiner t$_2$
with counted has there no-one

Moreover, I argued above that PP extraposition may induce the very same Anti-Freezing effect as remnant VP topicalization, because PP extraposition may be analyzed as adjunction to CP, i.e., as an instance of remnant movement. Cf. (70-a) and (70-b), which are also repeated here.

(90) a. dass da$_1$ keiner t$_2$ gerechnet hat $[_{PP}$ t$_1$ mit $]_2$
that there no-one counted has with

b. Wo$_1$ hat keiner t$_2$ gerechnet $[_{PP}$ t$_1$ mit $]_2$?
where has no-one counted with

Just like CPs, PPs can right-adjoin to a VP in German, provided that the VP undergoes topicalization. This is shown in (91):

(91) $[_{VP}$ $[_{VP}$ t$_2$ Gerechnet $]_3$ $[_{PP}$ damit $]_2$ $]_3$ hat der Fritz nicht t$_3$
counted there with has ART Fritz not

However, postposition stranding may never apply in such a configuration; consider (92), which is completely parallel to (78).

(92) *$[_{VP}$ $[_{VP}$ t$_2$ Gerechnet $]_3$ $[_{PP}$ t$_1$ mit $]_2$ $]_3$ hat da$_1$ der Fritz nicht t$_3$
counted with has there ART Fritz not

(92) involves scrambling of the NP *da*$_1$ to a pre-subject (IP-adjoined) position, extraposition of PP$_2$ to a VP-adjoined position, and topicalization of VP$_3$; and as with (78), it remains to be shown that there is no well-formed derivation of this example. If NP scrambling out of PP does not occur first, the Barriers Condition is violated. If NP scrambling to IP applies first, PP extraposition to VP applies second, and VP topicalization applies third, the second step violates the Strict Cycle Condition because PP extraposition has ended up in a lower position than the preceding scrambling operation did. Finally, if the order of rule application is *NP scrambling to IP* $\succ$ *VP topicalization* $\succ$ *PP extraposition*, the last step does not conform to the demands imposed by strict cyclicity. So far, so good. However, upon closer inspection things are a bit more complicated. Consider the following derivation of (92), which must also be excluded in order to rule out the sentence as ungrammatical (as before, the initial D-structure representation is simplified in various respects):

(93) a. *D-Structure Representation*:
– hat der Fritz nicht $[_{VP}$ $[_{PP}$ da$_1$ mit $]_2$ gerechnet $]_3$

b. *Scrambling of NP*$_1$ *to VP*:
– hat der Fritz nicht $[_{VP}$ da$_1$ $[_{VP}$ $[_{PP}$ t$_1$ mit $]_2$ gerechnet $]_3$ $]_3$

c. *Extraposition of PP*$_2$ *to VP*:
– hat der Fritz nicht $[_{VP}$ $[_{VP}$ da$_1$ $[_{VP}$ t$_2$ gerechnet $]_3$ $]_3$ $[_{PP}$ t$_1$ mit $]_2$ $]_3$

d. *Scrambling of NP*$_1$ *to IP*:
– hat da$_1$ der Fritz nicht $[_{VP}$ $[_{VP}$ t'_1 $[_{VP}$ t$_2$ gerechnet $]_3$ $]_3$ $[_{PP}$ t$_1$ mit $]_2$ $]_3$

e. *Topicalization of VP*$_3$:
$[_{VP}$ $[_{VP}$ t'_1 $[_{VP}$ t$_2$ gerechnet $]_3$ $]_3$ $[_{PP}$ t$_1$ mit $]_2$ $]_3$ hat da$_1$ der Fritz nicht t$_3$

This derivation does not violate the Barriers Condition (*da*-scrambling out of PP applies before PP extraposition) or the Strict Cycle Condition (PP extraposition ends up in the same minimal residue as the preceding scrambling operation).[35] Of course, this result can be achieved only by chain interleaving – scrambling of NP_1 is split up into two operations (adjunction to VP followed by adjunction to IP). Hence, we might hope that this derivation violates economy constraints. However, although this would indeed be the case under Collins' (1994) assumptions that I have abandoned, chain interleaving in (93) is not ruled out by Fewest Steps and Last Resort under present assumptions: Intermediate scrambling of NP_1 to VP is legitimate because the left-adjunction site of VP is a typical checking position for NP_1's [scr] feature (this satisfies Last Resort), but not yet (in this case) an actual one (this satisfies Fewest Steps).

The crucial problem with (93) seems to be that both the PP and the NP that it dominates are adjoined to VP from different directions (via extraposition and scrambling, respectively). I will assume that such a thing is ruled out by the following condition:

(94) If α dominates β, α-adjunction to γ and β-adjunction to γ cannot be bidirectional.

The basic idea behind (94) is that PP and NP in (93) compete for VP as an adjunction site, and that only one of the two items can be the winner of this competition. Thus, (94) rules out chain interleaving in the derivation in (93) and thereby closes the gap created by the option of intermediate scrambling here.[36] At present, the condition in (94) arguably has the character of brute force, and ideally, we would hope that it can be derived from deeper assumptions. I will show in chapter 5 (subsection 7.4.4) that (94) can indeed be made to follow from a revised version of Last Resort that incorporates Chomsky's (1995) Minimal Link Condition (MLC). However, for the time being, I will adopt (94) as a constraint on movement, and briefly explore some of its consequences.

First note that (94) does not preclude multiple left-adjunction or multiple right-adjunction. Furthermore, it does not rule out a combination of left- and right-adjunction to some XP if the two (or more) items that are adjoined do not stand in a dominance relation in the pre-movement structure. This is shown in (95-a) for multiple scrambling to IP, in (95-b) for multiple extraposition to VP (cf. Haider (1994, 3)

---

[35] An alternative derivation to (93) that poses exactly the same problem is one in which PP extraposition in the second step ends up in a lower VP-adjoined position than the preceding NP scrambling operation does. This derivation would not violate the Strict Cycle Condition either (since both landing sites are in the same minimal residue). In this derivation, VP topicalization in the last step could move only a lower VP segment, thereby circumventing an unbound intermediate trace t_1'. However, what is said about the derivation in (93) in what follows immediately carries over to this derivation, so we can abstract away from it for present purposes.

[36] Given this reasoning, there is independent theory-internal evidence for not assuming the split INFL hypothesis in German; recall footnote 15 of chapter 2. If AGR_OP and TP qualify as additional adjunction sites, a derivation should be possible that differs from the one in (93) in that PP extraposition ends up in a higher (AGR_OP-) right-adjoined position, and subsequent topicalization could involve AGR_OP rather than VP. Such a derivation would not violate any constraints under present assumptions.

for the basic observation), and in (95-c) for a combination of scrambling and extraposition to VP.

(95) a. dass [$_{\text{NP}}$ dem Fritz]$_1$ [$_{\text{NP}}$ das Buch]$_2$ keiner [$_{\text{VP}}$ t$_1$ t$_2$ geben]
that ART Fritz$_{dat}$ the book$_{acc}$ no-one$_{nom}$ give
wollte
wanted to

b. [$_{\text{VP}_3}$ [$_{\text{NP}}$ Einem Grammatiker t$_1$] t$_2$ aufgefallen [$_{\text{CP}_1}$ der das
a grammarian$_{dat}$ occurred to who that
untersuchte] [$_{\text{CP}_2}$ dass dieser Satz grammatisch ist]] ist erst letzte
investigated that this sentence grammatical is is only last
Woche t$_3$
week
'Only last week has it occurred to a grammarian who investigated it that this sentence is grammatical.'

c. [$_{\text{VP}_3}$ [$_{\text{PP}}$ Über die Liebe]$_1$ dieses Jahr [$_{\text{NP}}$ ein Buch t$_1$ t$_2$] gelesen [$_{\text{CP}_2}$
about the love this year a book read
das interessant war]] hat noch keiner t$_3$
that interesting was has yet no-one

Second, this approach makes an interesting prediction. Consider a minimal variant of (92) in which *da* ends up in a scrambling position below the landing site of the PP, so that it goes along with the VP under topicalization. This sentence is generated by a derivation that differs from the one in (93) only in that scrambling of *da* to IP does not take place (i.e., *da* checks the [scr] feature in a VP-adjoined position). Ceteris paribus, we expect the resulting sentence to be ill formed. The sentence is given in (96); it does indeed seem to be fairly deviant.

(96) ?*[$_{\text{VP}}$ [$_{\text{VP}}$ Da$_1$ [$_{\text{VP}}$ t$_2$ gerechnet]$_3$]$_3$ [$_{\text{PP}}$ t$_1$ mit]$_2$]$_3$ hat der Fritz nicht t$_3$
there counted with has ART Fritz not

A derivation of (96) might proceed via (a) scrambling of NP$_1$ to a left-adjoined VP position, followed by (b) extraposition of PP$_2$ to a higher right-adjoined VP position, followed by (c) topicalization of the VP that now contains a remnant PP which in turn contains an unbound trace t$_1$. This derivation does not violate the Barriers Condition, the Strict Cycle Condition, or Fewest Steps/Last Resort. However, since NP$_1$ is dominated by PP$_2$ in the pre-movement structure, it does violate the prohibition against bi-directional adjunction in (94).[37]

To sum up, it has turned out again that extraposition does not induce an Anti-Freezing effect per se: PP extraposition does not create an island if it ends up in a

[37] (96) is severely deviant, but it might not be ill formed to the same degree as (92). I have no explanation for a possible contrast in the acceptability of these sentences, though.

higher position than the preceding movement that extracts out of PP does, as in (90), but it creates an island if it ends up in a lower position, as in (92).[38]

5.2.3. Infinitive Extraposition to VP: The Third Construction

(97) instantiates the third construction in German:

(97) dass die Claudia das Buch$_1$ t$_2$ nicht versucht hat [$_\alpha$ dem Peter t$_1$ zu geben]$_2$
that ART Claudia the book$_{acc}$ not tried has ART Peter$_{dat}$ to give

I will assume along with den Besten & Rutten (1989) that examples like (97) involve a combination of (a) extraposition of an infinitival category α and (b) NP scrambling out of the infinitive.[39] Evidently, this construction exhibits an Anti-Freezing effect, and this is to be expected if α-extraposition can end up in a position that is higher than the one that the preceding NP scrambling operation ends up in (i.e., in an IP- or CP-right-adjoined position). Next, (98) shows that the infinitive can also be adjoined to VP, provided that VP undergoes topicalization:

(98) [$_{VP}$ t$_2$ Versucht [$_\alpha$ dem Peter das Buch$_1$ zu geben]$_2$]$_3$ hat die Claudia t$_3$ nicht
tried ART Peter$_{dat}$ the book$_{acc}$ to give has ART Claudia not

Now, consider (99), which is ungrammatical.

(99) *[$_{VP}$ t$_2$ Versucht [$_\alpha$ dem Peter t$_1$ zu geben]$_2$]$_3$ hat die Claudia das Buch$_1$ nicht t$_3$
tried ART Peter$_{dat}$ to give has ART Claudia the book$_{acc}$ not

Here, (a) NP scrambling out of the infinitive α and (b) topicalization of a VP in which α-extraposition has taken place occur in combination. Whereas these two operations are per se permitted (cf. (97) and (98), respectively), their combination is not; and

[38]There is a bit more to be said about constructions involving a combination of scrambling from PP and topicalization of VPs that contain unbound PP traces. I return to this issue in section 7 below.

[39]Den Besten & Rutten (1989) postulate that α = VP. As in the discussion of coherent vs. incoherent infinitives, I would like to remain uncommitted as to the categorial status of α here. The only thing that is relevant in the present context is that α does not block scrambling despite extraposition. Note in passing that the very existence of the third construction in German is not compatible with the standard view, according to which extraposition of an infinitive and transparency of the infinitive for scrambling are mutually exclusive (cf. Bech (1955/57), Stechow (1984), Grewendorf (1987), Sternefeld (1990), and Haider (1993), among others).

again, this follows from the analysis developed here – if α-extraposition to VP applies before NP scrambling, the Barriers Condition is violated;[40] if α-extraposition to VP applies after NP scrambling, the Strict Cycle Condition is violated (or condition (94), depending on assumptions about the position and projection of negation).[41]

Summarizing, it turns out that although "normal" cases of extraposition exhibit Anti-Freezing, extraposition may in fact induce Freezing effects under certain conditions – viz., if the landing site of extraposition is lower than the landing site of the preceding extraction operation. Given the interaction of the Barriers Condition, the Strict Cycle Condition (plus requirement (94)), and the combination of Fewest Steps and Last Resort (which rules out illicit chain interleaving), this is to be expected. From a more general point of view, it seems to me that the very existence of Freezing effects as they appear with extraposition to VP (cf. the ungrammatical examples in (78), (92), and (99)) poses a severe problem for analyses according to which V CP order in SOV languages is base-generated and not transformationally derived (recall the literature cited in subsection 4.2). Freezing effects clearly indicate movement, and I see no obvious way to accomodate the fact that V CP order creates an island for extraction in some contexts but not in others to an analysis in terms of base-generation of CP in its S-structure position.[42]

---

[40]This implies that left-adjunction to an extraposed infinitive is not a way to circumvent a barrier; cf. Grewendorf & Sabel (1994) for discussion.

[41]Sentences like (i) are much better than, e.g., (99) (Daniel Büring (p.c.)), and this is evidently a problem if (i) has essentially the same structure as (99), i.e., if it must be analyzed as indicated:

(i) ??[$_{\text{VP}}$ t$_2$ Versucht [$_\alpha$ t$_1$ zu schlagen]$_2$]$_3$ hat er den Mann$_1$ nicht t$_3$
tried to hit has he$_{nom}$ the man$_{acc}$ not

However, there is a crucial difference between (i) and (99) – in the former case (but not in the latter), the fronted VP is made up of two bare verb forms. I would like to suggest that the relative wellformedness of (i) is due to the (somewhat marginal) option of analyzing the sentence not via extraposition of an infinitival XP α that contains an argument trace, but rather via verb raising, i.e., head movement. According to this view, the structure of (i) actually looks as in (ii):

(ii) ??[$_{\text{VP}}$ [$_\alpha$ t$_1$ t$_V$]$_2$ [$_{\text{V}}$ Versucht [zu schlagen]$_V$]]$_3$ hat er den Mann$_1$ nicht t$_3$
tried to hit has he$_{nom}$ the man$_{acc}$ not

In all relevant respects, then, (ii) behaves like the most ordinary cases of remnant VP topicalization, because there is no instance of α-extraposition at all. Thus, a derivation of (ii) which proceeds in the order *verb raising* $\succ$ *NP scrambling* $\succ$ *VP topicalization* does not violate any constraint – in particular, extraction takes place only from complement positions, and every movement operation ends up in a higher position than the preceding movement operation does.

[42]Similar considerations apply with regard to analyses that view extraposition as a PF phenomenon; cf. footnote 21.

5.3. Extraposition in SVO Languages

Under the approach developed so far, we should also expect an Anti-Freezing effect with extraposition in SVO languages like English. And indeed, by and large, this prediction appears to be borne out. Thus, consider first an example involving a double object construction like (100).

(100) They explained$_1$ [$_{VP}$ [$_{NP}$ this] [$_{V'}$ t$_1$ [$_{PP}$ to the students]]]

Larson (1988) argues that the direct object *this* is base-generated in SpecV, and the indirect object *to the students* in the complement position of V. The verb then undergoes raising out of the VP into the head position of a VP-shell or some other functional projection. More generally, under this approach mapping of arguments of a verb into syntactic representations is regulated by a thematic hierarchy according to which a THEME argument (like *this* in (100)) always precedes and asymmetrically c-commands a GOAL argument (like *to the students* in (100)) at D-structure in English. If this is so, examples like the ones in (101) strongly suggest that clausal extraposition has taken place – note that in both cases, CP is the THEME argument of the verbs involved and hence must be located in a position higher than and to the left of the indirect object at D-structure.[43]

(101) a. They explained t$_2$ [$_{PP}$ to the students] [$_{CP}$ that they should read these books]$_2$
b. John promised t$_2$ [$_{PP}$ to Mary] [$_{CP}$ to leave]$_2$

Interestingly, the extraposition structures in (101) do not block *wh*-extraction; rather, an Anti-Freezing effect arises. This is shown by the examples in (102):

(102) a. What$_1$ did they explain t$_2$ [$_{PP}$ to the students] [$_{CP}$ that they should read t$_1$]$_2$?
b. What$_1$ did John promise t$_2$ [$_{PP}$ to Mary] [$_{CP}$ to do t$_1$]$_2$?

This Anti-Freezing effect is accounted for given that (102-a) and (102-b) each involve a derivation in which *wh*-movement of *what*$_1$ out of the direct object CP$_2$ in its in situ position is followed by right-adjunction of CP$_2$ to the matrix CP. Such a derivation ful-

---

[43] Actually, Larson (1988; 1991, 123) draws a different conclusion. He argues that what is involved in sentences like the ones at hand is not CP extraposition, but rather a process of "Light Predicate Raising," as shown in (i-ab).

(i) a. They [$_{V(')}$ explained to the students]$_1$ [$_{CP}$ that they should read these books]$_2$ t$_1$
b. John [$_{V(')}$ promised to Mary]$_1$ [$_{CP}$ to leave]$_2$ t$_1$

I will reject such an analysis because it involves either movement of a V' category or a non-structure-preserving rule of reanalysis which incorporates the indirect object into the verb so that a complex verb dominating an XP can undergo head movement to the left. Both options are precluded in the theory of movement adopted in chapter 2.

fills the demands imposed by the Strict Cycle Condition and the Barriers Condition.[44] If these considerations are on the right track, the conclusion to be drawn is that CP extraposition in English regularly exhibits Anti-Freezing effects, just like its German counterpart, and, consequently, that the English data are amenable to the same kind of analysis.[45]

6. Multiply Embedded CPs and Chain Interleaving

In this section, I want to address a consequence of the analysis that has been proposed here. This consequence concerns extraction from multiply embedded clauses in SOV languages like German, and it helps to differentiate between the two approaches to chain interleaving discussed above – i.e., Collins' (1994) strict prohibition against chain interleaving (derived from the standard version of Fewest Steps) on the one hand, and the weaker prohibition against chain interleaving (derived from the interaction of revised versions of Fewest Steps and Last Resort) developed in subsection 2.3.3 above. In the following subsection (6.1), I will begin by considering the second theory of chain interleaving, which I would like to argue for, and then turn to the predictions made by Collins' (1994) approach. Finally, in subsection 6.2, I will address

---

[44] A qualification is in order here. Given the notion of barrier that I have adopted so far (cf. (16)), the THEME CPs in (102) are in fact predicted to be barriers in their in situ positions (SpecV), contrary to fact. Thus, to accomodate the definition of barrier to VP-structures of the type proposed by Larson (1988; 1991), a modification is necessary. For present purposes, it may suffice to replace the notion of "complement position" in the original definition (16) by the more liberal notion of "selected position," where selected positions are either complement positions or VP-internal specifier positions:

(i) *Barrier* (revised):
XP is a barrier for α iff there is an X^n ($0 \leq n \leq P$) such that (a), (b), and (c) hold:
 a. X^n includes α.
 b. X^n is not selected.
 c. X^0 is distinct from Y^0, where XP is selected by Y^0.

See Müller & Sternefeld (1994; 1995) for further discussion.

[45] It seems, though, that the argument cannot be replicated with extraposition of NPs in English. In contrast to German, a combination of NP extraposition and *wh*-movement out of this NP does not seem to be possible in English; cf. (i) (from Lasnik & Saito (1992, 103); also see Wexler & Culicover (1980, 277ff)):

(i) ?*What$_1$ did you give t$_2$ to John [$_{NP}$ a book about t$_1$]$_2$?

The illformedness of (i) does not follow under the system of assumptions developed here. Lasnik & Saito (1992) suggest that the ungrammaticality of (i) does not result from a violation of the Barriers Condition, but "is due to the crossing effect" which is known to restrict movement in English (but not in German). I will assume that something along these lines is correct (i.e., that examples like (i) do exhibit an Anti-Freezing effect which is nevertheless masked by an intervening factor that creates ungrammaticality). As noted by João Costa (p.c.) and Daniel Büring (p.c.), similar considerations apply in the case of PP extraposition in English. Clearly, then, the "crossing effect" must be implemented in such a way that NPs and PPs are subject to it, whereas CPs are not.

some consequences of the less strict prohibition against chain interleaving that I adopt.

6.1. Extraction and Multiple Extraposition

Thus far, I have derived the Anti-Freezing effect that occurs with CP extraposition in sentences like (103), where the landing site of *wh*-movement is in the clause of which a segment immediately dominates the extraposed CP:

(103) (Ich weiß nicht) wen$_1$ er t$_2$ gesagt hat [$_{\text{CP}}$ t$'_1$ dass Claudia t$_1$ geküßt hat]$_2$
I know not whom he said has that Claudia kissed has

But now consider what the analysis developed here predicts for cases where the CP clause from which *wh*-movement takes place is more deeply embedded, as in (104).

(104) (Ich weiß nicht) [$_{\text{CP}_2}$ wen$_1$ er gesagt hat [$_{\text{CP}_3}$ dass Peter denkt [$_{\text{CP}_4}$ dass
I know not whom he said has that Peter thinks that
Claudia t$_1$ geküßt hat]]]
Claudia kissed has

Given that V CP order in SOV languages can be derived only by extraposition, as I have assumed throughout, the D-structure representation of (104) must look roughly as in (105).

(105) ... [$_{\text{CP}_2}$ – er [$_{\text{CP}_3}$ – dass Peter [$_{\text{CP}_4}$ – dass Claudia wen$_1$ geküßt hat] denkt] gesagt hat]

The task now is to show that there is a well-formed derivation of sentence (104) on the basis of (105). In order to escape from the most deeply embedded clause CP$_4$, *wh*-movement of *wen*$_1$ (‘whom$_{acc}$’) must take place prior to CP$_4$ extraposition; otherwise, *wh*-movement will violate the Barriers Condition (a right-adjoined CP$_4$ is not in a selected position). Hence, we can disregard from the start any derivation in which *wh*-movement applying to *wen*$_1$ is not the first movement step. The question then is, Where does initial *wh*-movement of *wen*$_1$ end up?

A first possibility is that the *wh*-phrase *wen* is moved to the matrix SpecC position (i.e., SpecC$_2$) first, via three instances of *wh*-movement. Since all CPs are still in situ, locality requirements (in particular, the Barriers Condition) are met. Suppose further that the next movement operation is right-adjunction of CP$_4$ to CP$_3$, followed finally by CP$_3$ extraposition to CP$_2$. In this derivation, CP$_4$ extraposition in the second to last step violates the Strict Cycle Condition: Because of *wh*-movement to SpecC of CP$_2$, CP$_2$ is a cyclic node for CP$_4$ extraposition, but this movement affects only CP$_3$, and not CP$_2$, as would be required. Therefore, the derivation in (106), although compatible with locality constraints and economy constraints (neither Fewest Steps nor Last Resort is violated), is in conflict with strict cyclicity.

(106) a. *Wh-Movement of ‘wen’ to SpecC of CP$_4$*:
... [$_{\text{CP}_2}$ – er [$_{\text{CP}_3}$ – dass Peter [$_{\text{CP}_4}$ wen$_1$ dass Claudia t$_1$ geküßt hat] denkt

] gesagt hat]

b. *Wh-Movement of 'wen' to SpecC of CP$_3$*:
... [$_{CP_2}$ – er [$_{CP_3}$ wen$_1$ dass Peter [$_{CP_4}$ t'_1 dass Claudia t_1 geküßt hat] denkt] gesagt hat]

c. *Wh-Movement of 'wen' to SpecC of CP$_2$*:
[$_{CP_2}$ wen$_1$ er [$_{CP_3}$ t''_1 dass Peter [$_{CP_4}$ t'_1 dass Claudia t_1 geküßt hat] denkt] gesagt hat]

d. *Extraposition of CP$_4$ to CP$_3$*:
[$_{CP_2}$ wen$_1$ er [$_{CP_3}$ [$_{CP_3}$ t''_1 dass Peter t_4 denkt] [$_{CP_4}$ t'_1 dass Claudia t_1 geküßt hat]] gesagt hat]

e. *Extraposition of CP$_3$ to CP$_2$*:
[$_{CP_2}$ [$_{CP_2}$ wen$_1$ er t_3 gesagt hat] [$_{CP_3}$ [$_{CP_3}$ t''_1 dass Peter t_4 denkt] [$_{CP_4}$ t'_1 dass Claudia t_1 geküßt hat]]]

Since (106-d) is the illicit movement step in (106) (CP$_2$ extraposition violates the Strict Cycle Condition), one might try to reverse the two CP extraposition operations, as in the following derivation:

(107) a. *Wh-Movement of 'wen' to SpecC of CP$_4$*:
... [$_{CP_2}$ – er [$_{CP_3}$ – dass Peter [$_{CP_4}$ wen$_1$ dass Claudia t_1 geküßt hat] denkt] gesagt hat]

b. *Wh-Movement of 'wen' to SpecC of CP$_3$*:
... [$_{CP_2}$ – er [$_{CP_3}$ wen$_1$ dass Peter [$_{CP_4}$ t'_1 dass Claudia t_1 geküßt hat] denkt] gesagt hat]

c. *Wh-Movement of 'wen' to SpecC of CP$_2$*:
[$_{CP_2}$ wen$_1$ er [$_{CP_3}$ t''_1 dass Peter [$_{CP_4}$ t'_1 dass Claudia t_1 geküßt hat] denkt] gesagt hat]

d. *Extraposition of CP$_3$ to CP$_2$*:
[$_{CP_2}$ [$_{CP_2}$ wen$_1$ er t_3 gesagt hat] [$_{CP_3}$ t''_1 dass Peter [$_{CP_4}$ t'_1 dass Claudia t_1 geküßt hat] denkt]]

e. *Extraposition of CP$_4$ to CP$_3$*:
[$_{CP_2}$ [$_{CP_2}$ wen$_1$ er t_3 gesagt hat] [$_{CP_3}$ [$_{CP_3}$ t''_1 dass Peter t_4 denkt]] [$_{CP_4}$ t'_1 dass Claudia t_1 geküßt hat]]

Again, all three *wh*-movement operations required by locality constraints apply first; this turns CP$_2$ into a cyclic node. Next, however, CP$_3$ (and not CP$_4$) undergoes extraposition (to CP$_2$). This step does not violate strict cyclicity because it ends up in the minimal residue of the head of CP$_2$. The final step, however, right-adjoins CP$_4$ to CP$_3$. This step violates the Strict Cycle Condition because an adjunction site of an XP (CP$_3$, in this case) that is itself adjoined to another XP (CP$_2$) is not part of the minimal residue of the head of this latter XP; recall that under present assumptions a non-recursive notion of minimal residue must be adopted, as in (108) (= (84)):

(108) *Minimal Residue*:
The minimal residue of a head X comprises X-Adj, SpecX, and XP-Adj, and nothing else.

Thus, extraposition of CP_4 in this derivation ends up in a position that is dominated by CP_2 but does not belong to the minimal residue of C_2, and the derivation is ill formed.

As a third possibility, consider the derivation in (109).

(109) a. *Wh-Movement of 'wen' to SpecC of CP_4*:
... [$_{CP_2}$ – er [$_{CP_3}$ – dass Peter [$_{CP_4}$ wen$_1$ dass Claudia t_1 geküßt hat] denkt] gesagt hat]

b. *Wh-Movement of 'wen' to SpecC of CP_3*:
... [$_{CP_2}$ – er [$_{CP_3}$ wen$_1$ dass Peter [$_{CP_4}$ t_1' dass Claudia t_1 geküßt hat] denkt] gesagt hat]

c. *Wh-Movement of 'wen' to SpecC of CP_2*:
[$_{CP_2}$ wen$_1$ er [$_{CP_3}$ t_1'' dass Peter [$_{CP_4}$ t_1' dass Claudia t_1 geküßt hat] denkt] gesagt hat]

d. *Extraposition of CP_4 to CP_2*:
[$_{CP_2}$ [$_{CP_2}$ wen$_1$ er [$_{CP_3}$ t_1'' dass Peter t_4 denkt] gesagt hat] [$_{CP_4}$ t_1' dass Claudia t_1 geküßt hat]]

e. *Extraposition of CP_3 to CP_2*:
[$_{CP_2}$ [$_{CP_2}$ [$_{CP2}$ wen$_1$ er gesagt hat] [$_{CP_3}$ t_1'' dass Peter t_4 denkt]] [$_{CP_4}$ t_1' dass Claudia t_1 geküßt hat]]

As in the derivations (106) and (107), *wh*-movement takes place to the ultimate landing site first, i.e., to the SpecC position of CP_2. Next, as in (106), CP_4 extraposition applies. However, this time, extraposition of CP_4 does not go to a right-adjunction site of CP_3, which would violate strict cyclicity, but rather to a right-adjunction site of CP_2, in accordance with strict cyclicity; finally, CP_3 extraposition to CP_2 applies. What is wrong with this derivation?[46] I think that what is wrong is that the movement operation in (109-d) (CP_4 extraposition to CP_2) involves non-clause-bound extraposition of CP_4 (skipping over CP_3), and this seems to be precluded on general grounds. Long-distance extraposition is commonly excluded by the Upward Boundedness Constraint (or Right Roof Constraint) that goes back to Ross (1967). For present purposes, the descriptive statement in (110) may suffice; for attempts to derive it from independently motivated constraints, see, e.g., Culicover & Rochemont (1990) and Müller &

[46]Note that extraposition of CP_3 to CP_2 in the last step in (109) does not violate strict cyclicity. The reason is that although CP_2 is a cyclic node for extraposition of CP_3 and although this movement step ends up in a position that is lower than that of CP_4, the landing site of CP_3 is still part of the minimal residue of the head of CP_2, as required by the Strict Cycle Condition.

Sternefeld (1993).[47]

(110) *Upward Boundedness Constraint*:
Rightward movement must not cross a CP.

To sum up so far, it seems that the derivations in (106), (107), and (109) cannot generate the sentence in (104) without violating constraints. A common feature of all these derivations is that all *wh*-movement operations apply before all extraposition operations. I would like to contend that this view is in fact not correct and that a legitimate derivation of the string in (104) must take the following form:

(111) a. *Wh-Movement of 'wen' to SpecC of* CP_4:
... [$_{CP_2}$ – er [$_{CP_3}$ – dass Peter [$_{CP_4}$ wen$_1$ dass Claudia t$_1$ geküßt hat] denkt] gesagt hat]

b. *Wh-Movement of 'wen' to SpecC of* CP_3:
... [$_{CP_2}$ – er [$_{CP_3}$ wen$_1$ dass Peter [$_{CP_4}$ t_1' dass Claudia t$_1$ geküßt hat] denkt] gesagt hat]

c. *Extraposition of* CP_4 *to* CP_3:
[$_{CP_2}$ – er [$_{CP_3}$ [$_{CP_3}$ wen$_1$ dass Peter t$_4$ denkt] [$_{CP_4}$ t_1' dass Claudia t$_1$ geküßt hat]] gesagt hat]

d. *Wh-Movement of 'wen' to SpecC of* CP_2:
[$_{CP_2}$ wen$_1$ er [$_{CP_3}$ [$_{CP_3}$ t_1'' dass Peter t$_4$ denkt] [$_{CP_4}$ t_1' dass Claudia t$_1$ geküßt hat]] gesagt hat]

e. *Extraposition of* CP_3 *to* CP_2:
[$_{CP_2}$ [$_{CP_2}$ wen$_1$ er t$_3$ gesagt hat] [$_{CP_3}$ [$_{CP_3}$ t_1'' dass Peter t$_4$ denkt] [$_{CP_4}$ t_1' dass Claudia t$_1$ geküßt hat]]]

Here, *wh*-movement out of a CP precedes extraposition of that CP, in accordance with the Barriers Condition. Furthermore, each CP extraposition operation ends up in the minimal residue of the highest cyclic node (that has been created by preceding *wh*-movement), in accordance with the Strict Cycle Condition. In addition, the Upward Boundedness Constraint is respected because CP extraposition never skips over an intervening CP. Thus, this derivation exhibits a "seesaw" effect, in the sense that *wh*-extraction and CP extraposition alternate systematically. Of course, this seesaw derivation involves chain interleaving, and the question arises of whether this is permitted.

---

[47] The Upward Boundedness Constraint is also responsible for the fact that sentences like (i) are ungrammatical:

(i) *Ich weiß nicht [$_{CP_2}$ wen er gesagt hat [$_{CP_4}$ dass Claudia geküßt hat] [$_{CP_3}$ dass Peter denkt]]
I know not whom he said has that Claudia kissed has that Peter thinks

(i) can only be derived via a derivation involving long-distance extraposition of CP_4. Note that the only difference between this derivation generating (i) and the one in (109) is that CP_3 extraposition in the final step ends up in a *lower* CP_2-adjoined position than the preceding operation applying to CP_4 in (109), and in a *higher* CP_2-adjoined position in (i).

Under the revised notions of Fewest Steps and Last Resort developed in chapter 2 and adopted here, it turns out that it is. The reason is that although there are several movement operations that apply to the *wh*-phrase, they are all of the same type (viz., *wh*-movement operations to a SpecC substitution position) and therefore a PUB effect, reducible to the interaction of Fewest Steps and Last Resort, does not arise.

With all this in mind, consider now the prediction that the approach to chain interleaving developed by Collins (1994) makes for sentences like (104). Recall that Collins (1994) adopts the standard definition of the Fewest Steps condition, according to which a derivation is blocked by a competing derivation if the latter one involves fewer operations; and recall that this implies that chain interleaving is generally precluded and that successive-cyclic movement must be handled in terms of the notion of Form Chain. From this it follows that a derivation like (111) can never be legitimate, because it crucially involves chain interleaving (applying to the *wh*-phrase) of a kind that cannot be avoided (i.e., explained away) by invoking Form Chain. The derivations in (106), (107), and (109), on the other hand, do not violate the standard Fewest Steps condition, assuming that the first three instances of *wh*-movement are conceived of as one Form Chain operation. However, we have just seen that these three derivations violate either the Strict Cycle Condition or the Upward Boundedness Constraint (other derivations in which all *wh*-movement is postponed to the end violate the Barriers Condition); the only candidate for a legitimate derivation is (111), which involves chain interleaving. Hence, under the present system of assumptions, we have yet another argument in favour of replacing the standard notions of Fewest Steps and Last Resort with those in (35) and (36) (repeated here in (112) and (113)), in addition to those arguments given in section 6 of chapter 2.

(112) *Fewest Steps*:
If two derivations D_1 and D_2 are in the same reference set and D_1 involves fewer checking operations than D_2, then D_1 is to be preferred over D_2.

(113) *Last Resort*:
α is raised to a position β only if β is a typical checking position for the lowest-ranked unchecked morphological feature of α.

More generally, we have (theory-internal) evidence that chain interleaving is not to be excluded on principled grounds (rather, it is ruled out only if it involves a PUB effect, reducible to the interaction of Fewest Steps and Last Resort) and that the concept of Form Chain not only is superfluous and conceptually problematic (as argued in subsection 5.1 of chapter 2), but actually makes wrong predictions in the case of *wh*-extraction from multiply embedded CPs in SOV languages with CP extraposition.

6.2. Other Cases of Chain Interleaving

Of course, the present, more variable approach to chain interleaving in terms of the revised notions of Fewest Steps and Last Resort differs in a number of other respects from Collins' (1994) strict approach that relies on the standard notion of Fewest Steps. In general, there are now several cases of – apparently illicit – chain interleaving that are no longer ruled out by economy. In most of these cases, it seems to me that an alternative explanation of the illformedness of the resulting derivation is readily available; but some cases prove to be a little bit more complicated.

By way of illustration, consider the following schematic derivation via chain interleaving of an ungrammatical example involving a Freezing effect (as before, linear ordering is irrelevant as such, but supposed to correspond to asymmetric c-command):

(114) a. *D-Structure Representation*:
... $[_\beta$ ... α_1 ... $]_2$
b. α-*Movement no.1*:
... α_1... $[_\beta$ ... t_1 ... $]_2$
c. β-*Movement*:
... $[_\beta$ ... t_1 ... $]_{2...}$ α_1... t_2
d. α-*Movement no.2*:
... α_1 ... $[_\beta$ ... t_1 ... $]_{2...}$ t'_1... t_2

Suppose for the sake of the argument that α-movement in the first and third steps ends up in the same kind of position, so that a PUB effect is avoided (i.e., Fewest Steps and Last Resort are fulfilled). Nevertheless, movement in (114) might be illegitimate. For instance, α-movement in (114-b) might somehow fill an escape hatch or create a strict island (via the Barriers Condition) which then blocks β-movement in the following step. Similarly, β-movement in (114-c) might block an escape hatch or create a strict island for subsequent α-movement in (114-d).[48] But again, suppose that neither α-movement nor β-movement in (114) creates a strict island and bars further extraction due to an unavoidable violation of some locality constraint. Among the movement types that seem to never create islands are left-adjunction (i.e., scrambling) operations (see Müller & Sternefeld (1993) vs. Lasnik & Saito (1992)) and right-adjunction (i.e., extraposition) operations. The property of not creating islands that is found with scrambling and extraposition is, of course, crucial for the very existence of Anti-Freezing effects of the type discussed up to now; otherwise, a sentence like, e.g., (115) that is derived by remnant VP topicalization following NP scrambling could never be well-formed:

(115) $[_{CP}$ $[_{VP}$ t_1 Gelesen $]_2$ hat $[_{NP}$ das Buch $]_1$ keiner t_2 $]$
read has the book no-one

Thus, if all movement operations in (114) are instances of scrambling, we do not anticipate a violation of economy (Fewest Steps & Last Resort), locality (Barriers

[48] These island-creating properties of certain movement types are well known; see, e.g., the strict islandhood imposed on a clause by internal topicalization in the Germanic languages, as discussed by den Besten (1989), Rochemont (1989), and Vikner (1995), among many others.

Condition), or cyclicity (Strict Cycle Condition – note that all movement ends up in a higher position than any preceding movement operation in (114)). Recall now from subsection 2.1.2 that scrambling of an XP creates a Freezing effect if XP contains a bound trace at S-structure. Above, I gave examples involving bound *wh*-traces in scrambled items, such as (7-b), which is repeated here as (116):

(116) *Wo$_1$ meinst du [$_{CP}$ t$'_1$ dass [$_{PP}$ t$_1$ mit]$_2$ keiner t$_2$ gerechnet hat] ?
what think you that with no-one counted has

However, the following examples show that scrambling also creates a Freezing effect in the case of bound scrambling traces:[49]

(117) a. *dass [$_{NP}$ den Aufsatz]$_1$ mal wieder [$_{VP}$ t$_1$ gelesen]$_2$ keiner t$'_1$ t$_2$ hat
that the article$_{acc}$ once again read no-one$_{nom}$ has

b. *dass darüber$_1$ gestern [$_{NP}$ ein Buch t$_1$]$_2$ keiner t$'_1$ t$_2$ gelesen hat
that about that yesterday a book$_{acc}$ no-one$_{nom}$ read has

c. *dass da$_1$ manchmal [$_{PP}$ t$_1$ mit]$_2$ keiner t$'_1$ t$_2$ rechnen kann
that there sometimes with no-one$_{nom}$ count can

A derivation of, e.g., (117-a) via chain interleaving might look as in (118).

(118) a. *D-Structure Representation*:
dass mal wieder keiner [$_{VP}$ [$_{NP}$ den Aufsatz]$_1$ gelesen]$_2$ hat

b. *NP$_1$ Scrambling to VP*:
dass mal wieder keiner [$_{NP}$ den Aufsatz]$_1$ [$_{VP}$ t$_1$ gelesen]$_2$ hat

c. *VP$_2$ Scrambling to IP*:
dass mal wieder [$_{VP}$ t$_1$ gelesen]$_2$ keiner [$_{NP}$ den Aufsatz]$_1$ t$_2$ hat

d. *NP$_1$ Scrambling to IP*:
dass [$_{NP}$ den Aufsatz]$_1$ mal wieder [$_{VP}$ t$_1$ gelesen]$_2$ keiner t$'_1$ t$_2$ hat

As it stands, the derivation in (118) does not violate any constraint. Locality requirements are met (extraction out of an XP takes place only if the XP is in situ in a complement – or selected – position), strict cyclicity is respected, and chain interleaving is permitted under present assumptions: The two instances of NP$_1$ scrambling end up in positions of the same type (i.e., left-adjoined positions), so both movement operations applying to NP$_1$ are permitted by Last Resort/Fewest Steps. Thus, at first sight it seems as though excluding a derivation like that in (118) is a problem in the

[49](117-a) involves VP scrambling and would therefore have a marked status irrespective of the issue of Freezing; see footnote 14.

present approach – this derivation involves chain interleaving, but of a kind that is apparently permitted, in complete analogy to the *wh*-movement case (111) (or any other instance of successive-cyclic *wh*-movement, for that matter).

As a first step towards a possible solution, consider again an example like (43), repeated here as (119):

(119) *weil [$_{VP}$ t$_1$ gelesen]$_2$ [$_{NP}$ das Buch]$_1$ keiner t$_2$ hat
because read the book no-one has

In subsection 3.1, I took this example to illustrate the illegitimacy of remnant scrambling in cases where the antecedent of the unbound trace has also undergone scrambling; cf. (120-a). A slightly more general way of stating this would be (120-b).

(120) a. There are no unbound scrambling traces in scrambled items.
b. There are no scrambling traces in scrambled items.

Clearly, if the descriptive generalization in (120-b) can be derived, both illicit remnant scrambling in (119) and the illicit cases of scrambling in (117) (with derivations as in (118)) are accounted for. If, on the other hand, only the generalization in (120-a) can be derived from independent assumptions, it seems as though the problem with the data in (117) persists. For the time being, let us assume that (120-b) does indeed follow in one way or another (I will address this issue in some detail in chapter 5). If so, there is no problem with (117) anymore, even though chain interleaving in a derivation like (118) is not ruled out as such.

There are other cases of apparently illicit chain interleaving that take the general form of (114); but I would like to contend that they either involve a violation of locality, economy or cyclicity constraints as discussed so far or can be explained away, as in the case of (118).

Concluding this subsection, it has proven possible to account for illicit cases of chain interleaving in a derivation by the interaction of the revised versions of Fewest Steps and Last Resort argued for in chapter 2. The interaction of these constraints, however, is not fully equivalent to the approach pursued by Collins (1994) that rests on the standard notion of Fewest Steps. In particular, certain cases of chain interleaving that are precluded under the standard approach are now allowed; and I have tried to argue that this result is indeed a desirable one, given data such as (104) that involve successive-cyclic *wh*-movement from deeply embedded clauses in German.

Much more could be said about the interaction of extraposition and extraction in the present framework. For the time being, though, I will highlight only one further consequence that the analysis might have; I will argue that an otherwise unexpected Anti-Freezing effect can be accounted for by invoking the idea that invisible extraposition is involved.[50]

7. (Anti-) Freezing and String-Vacuous Scrambling

7.1. The Problem: Unexpected Anti-Freezing

Under the remnant movement approach, PP_2 in (121-a) and NP_2 in (121-b) occupy (IP-) left-adjoined positions as a result of a preceding scrambling operation that has removed these items from the VP:

(121) a. [$_{VP}$ t_2 Gerechnet $]_3$ hat wie immer [$_{PP}$ damit $]_2$ keiner
counted has as always therewith no-one

b. [$_{VP}$ t_2 Gelesen $]_3$ hat gestern [$_{NP}$ ein Buch darüber $]_2$ der Fritz
read has yesterday a book$_{acc}$ about that ART Fritz$_{nom}$

VP topicalization produces an Anti-Freezing effect, for reasons discussed at length. However, given what has just been said about the derivation in (118), we would expect a Freezing effect to show up if the sentences in (121) were to involve an additional scrambling operation that removes an item from the scrambled XP_2 (the R-pronoun *da* and the PP *darüber*, respectively). As shown in (122), this prediction is borne out:

(122) a. *[$_{VP}$ t_2 Gerechnet $]_3$ hat da_1 wie immer [$_{PP}$ t_1 mit $]_2$ keiner t_3
counted has there as always with no-one

b. *[$_{VP}$ t_2 Gelesen $]_3$ hat $darüber_1$ gestern [$_{NP}$ ein Buch t_1 $]_2$ der Fritz t_3
read has about that yesterday a book$_{acc}$ ART Fritz$_{nom}$

Now, consider the examples in (123), which differ from their counterparts in (121) only insofar as PP_2 and NP_2 have undergone *string-vacuous* scrambling, according to present assumptions:

[50] Another interesting consequence of the present analysis of extraposition is related to binding theory. At first sight, it might look as though the hypothesis that right-adjunction can always go to any kind of XP faces empirical problems. Thus, Baltin (1983), Reinhart (1983), and Culicover & Rochemont (1990) claim that extraposition has binding-theoretic effects in English and that there is a difference between extraposition to VP and extraposition to IP in this respect. (Also see Fanselow (1987) and Grewendorf (1988) for the same kind of argument with respect to German.) Such a view is not compatible with the assumption that extraposition of some item α must end in a CP-adjoined position if, e.g., *wh*-movement out of α to a higher SpecC position is to take place. However, upon closer inspection it appears that the evidence from binding theory is by far not as decisive as one might assume (cf., e.g., the contrast between Reinhart's and Culicover & Rochemont's assumptions); moreover, as observed by Büring & Hartmann (1997), there is reason to believe that the base position is much more important than the derived (extraposition) position for the evaluation of binding options. For reasons of space and coherence, though, I will abstract away from these issues in what follows.

(123) a. [$_{VP}$ t$_2$ Gerechnet]$_3$ hat wie immer keiner [$_{PP}$ damit]$_2$ t$_3$
counted has as always no-one therewith

b. [$_{VP}$ t$_2$ Gelesen]$_3$ hat gestern der Fritz [$_{NP}$ ein Buch
read has yesterday ART Fritz$_{nom}$ a book$_{acc}$
darüber]$_2$ t$_3$
about that

If PP$_2$ and NP$_2$ occupy scrambling positions in these examples, too, then we clearly expect them to create Freezing effects if further scrambling of the R-pronoun *da* or the PP *darüber* takes place. However, as noted by den Besten & Webelhuth (1990, 87), this is not the case; the relevant sentences are well formed, in contrast to their counterparts in (122); cf.:

(124) a. [$_{VP}$ t$_2$ Gerechnet]$_3$ hat da$_1$ wie immer keiner [$_{PP}$ t$_1$ mit]$_2$ t$_3$
counted has there as always no-one with

b. [$_{VP}$ t$_2$ Gelesen]$_3$ hat darüber$_1$ gestern der Fritz [$_{NP}$ ein
read has about that yesterday ART Fritz$_{nom}$ a
Buch t$_1$]$_2$ t$_3$
book$_{acc}$

Before I try to tackle this problem, I would like to briefly focus on another peculiar property related to string-vacuous scrambling in remnant movement contexts.

7.2. Categorial Selectivity

It has often been observed (cf. Webelhuth (1987), Stechow & Sternefeld (1988), Frey & Tappe (1991), Haider (1993), and Müller (1995), among others) that scrambling in German is categorially selective in the sense that certain kinds of categories do not scramble easily. This holds, e.g., for predicates in general (see chapter 5 below). It also holds for indefinite pronouns that take the form of *wh*-pronouns. The restriction is shown in (125-a) for predicates and in (125-b) for indefinite *wh*-pronouns:

(125) a. ??dass [$_{AP}$ roh]$_1$ der Karl das Fleisch t$_1$ gegessen hat
that raw ART Karl$_{nom}$ the meat$_{acc}$ eaten has

b. ??dass Antje was$_1$ gestern t$_1$ gelesen hat
that Antje$_{nom}$ something yesterday read has

For present purposes, it is immaterial why scrambling does not lead to a perfectly acceptable result in these contexts. The important thing is that ceteris paribus, we expect exactly the same intermediate status to arise in the case of remnant VP topicalization, where the antecedent of the unbound trace is in a scrambling position. If such scrambling is not string-vacuous, this prediction is borne out, as (126) shows:

(126) a. ??[$_{VP}$ t$_1$ Gegessen]$_2$ hat der Karl [$_{AP}$ roh]$_1$ das Fleisch t$_2$
eaten has ART Karl$_{nom}$ raw the meat$_{acc}$

b. ??[$_{VP}$ t$_1$ Gelesen]$_2$ hat Antje was$_1$ gestern t$_2$
read has Antje something yesterday

However, as noted by Haider (1991, 29) and Jochen Geilfuß (p.c.), if scrambling is string-vacuous, the weak ban on scrambling of indefinite pronouns that have the form of *wh*-pronouns is lifted; cf. (127-b). It seems to me that the same goes for the weak prohibition against predicate scrambling; cf. (127-a):

(127) a. [$_{VP}$ t$_1$ Gegessen]$_2$ hat der Karl das Fleisch [$_{AP}$ roh]$_1$ t$_2$
eaten has ART Karl$_{nom}$ the meat$_{acc}$ raw

b. [$_{VP}$ t$_1$ Gelesen]$_2$ hat Antje gestern was$_1$ t$_2$
read has Antje yesterday something

To sum up so far, string-vacuous scrambling in remnant movement configurations fails to exhibit two properties that are usually associated with scrambling in German – it does not create a Freezing effect (even though it is not remnant movement by itself, i.e., even though the scrambled item does not contain an unbound trace), and it does not show categorial selectivity. It seems plausible to relate these two properties. In what follows, I will try to sketch a tentative solution to this problem.

7.3. String-Vacuous Scrambling as Extraposition

If the remnant movement approach is to be maintained in light of examples such as those in (124) (which are repeated here for convenience in (128)), it seems that there are two possible ways out.

(128) a. [$_{VP}$ t$_2$ Gerechnet]$_3$ hat da$_1$ wie immer keiner [$_{PP}$ t$_1$ mit]$_2$ t$_3$
counted has there as always no-one with

b. [$_{VP}$ t$_2$ Gelesen]$_3$ hat darüber$_1$ gestern der Fritz [$_{NP}$ ein Buch t$_1$]$_2$ t$_3$
read has about that yesterday ART Fritz$_{nom}$ a book$_{acc}$

On the one hand, one might want to modify the theory of barriers and locality, to the effect that XP$_2$ in (128) is not an island even though it occupies a left-adjoined scrambling position.[51] On the other hand, one might try to show that string-vacuous scrambling is not the only option available for PP$_2$ and NP$_2$ in (128); rather, it might be the case that these items have undergone some other kind of movement that exempts them from inducing a Freezing effect. It is this latter strategy that I would like to

---

[51] Such an approach is pursued by den Besten & Webelhuth (1990, 87-91), who suggest that either scrambling does not create a barrier if it is highly local or, alternatively, all instances of remnant movement are reconstructed at LF and S-structure extraction is classified as legitimate if adjacency of XP$_2$ and the head of a reconstructed VP can be accomplished at LF.

pursue here.

Note first that PP_2 and NP_2 in (128) occupy right-peripheral positions in their respective clauses. Hence, it could in principle be possible that these items are extraposed. Suppose that this is so and that the landing site is a CP-right-adjoined position. If that is so, the Anti-Freezing effect with movement of PP_2 and NP_2 in (128) is explained. (128-a), e.g., can have the following derivation:

(129) a. *D-Structure Representation*:
– hat wie immer keiner $[_{\text{VP}}$ $[_{\text{PP}}$ da_1 mit $]_2$ gerechnet $]_3$
b. *Scrambling of NP_1 to IP*:
– hat da_1 wie immer keiner $[_{\text{VP}}$ $[_{\text{PP}}$ t_1 mit $]_2$ gerechnet $]_3$
c. *Extraposition of PP_2 to CP*:
– hat da_1 wie immer keiner $[_{\text{VP}}$ t_2 gerechnet $]_3$ $[_{\text{PP}}$ t_1 mit $]_2$
d. *Topicalization of VP_3*:
$[_{\text{VP}}$ t_2 gerechnet $]_3$ hat da_1 wie immer keiner t_3 $[_{\text{PP}}$ t_1 mit $]_2$

This derivation respects all the locality, economy, and cyclicity constraints adopted so far: Extraction of NP_1 from PP_2 takes place while PP_2 is still in situ (so the Barriers Condition is fulfilled); extraposition of PP_2 turns CP into a cyclic node, but subsequent VP topicalization ends up in the same minimal residue, in accordance with the Strict Cycle Condition.[52]

If this analysis is on the right track, we can expect that minimally different cases in which we can see that there is no ambiguity between scrambling and extraposition, with only the former option being possible, should result in illformedness. Although the data are quite murky here and judgements are not solid, it seems to me that this prediction is borne out, at least as a strong tendency. To see this, consider first the sentences in (130), which are analogous to those in (128) and well formed:

(130) a. $[_{\text{VP}}$ t_2 Rechnen können $]_3$ hat da_1 keiner $[_{\text{PP}}$ t_1 mit $]_2$
count can has there no-one_{nom} with
b. $[_{\text{VP}}$ t_2 Lesen wollen $]_3$ hat darüber_1 keiner $[_{\text{NP}}$ ein Buch t_1 $]_2$
read want to has about that no-one_{nom} a book_{acc}

Here, the topicalized VP is headed by a modal verb that embeds the VP headed by the lexical verb; this latter VP contains an unbound trace of XP_2. Nothing speaks against assuming that PP_2 and NP_2 are extraposed, and this accounts for the wellformedness of the sentences. Of course, the VPs headed by the lexical verbs can also undergo topicalization alone, thereby creating an unbound trace of XP_2, as shown in (131):

[52] For the sake of simplicity, I abstract away from the distinction between CP and TopP here. Since TopP is dominated by CP, PP extraposition in the second step must actually end up in a TopP-adjoined position, if nothing more is said. See chapter 6 for related remarks.

(131) a. [$_{VP}$ t$_2$ Rechnen]$_4$ hat keiner [$_{PP}$ damit]$_2$ [$_{VP}$ t$_4$ können]$_3$
count has no-one$_{nom}$ therewith can

b. [$_{VP}$ t$_2$ Lesen]$_4$ hat keiner [$_{NP}$ ein Buch darüber]$_2$ [$_{VP}$ t$_4$
read has no-one$_{nom}$ a book$_{acc}$ about that
wollen]$_3$
want to

But note that extraction from XP$_2$ has not taken place in (131). As soon as such extraction occurs, the resulting sentence is significantly degraded. Compare the sentences in (130) with those in (132):

(132) a. ??[$_{VP}$ t$_2$ Rechnen]$_4$ hat da$_1$ keiner [$_{PP}$ t$_1$ mit]$_2$ [$_{VP}$ t$_4$ können]$_3$
count has there no-one$_{nom}$ with can

b. ?*[$_{VP}$ t$_2$ Lesen]$_4$ hat darüber$_1$ keiner [$_{NP}$ ein Buch t$_1$]$_2$ [$_{VP}$ t$_4$
read has about that no-one$_{nom}$ a book$_{acc}$
wollen]$_3$
want to

This conforms to expectations: If PP$_2$ and NP$_2$ are scrambled, this will invariably result in a Freezing effect; and if these items are extraposed, extraposition cannot possibly have ended up in a CP-adjoined position due to the intervening embedding verb that is the head of VP$_3$. Extraposition could only have ended up in a right-adjoined position of VP$_4$ (with the lower VP$_4$ segment subsequently undergoing topicalization); but in that case, a violation of the Strict Cycle Condition will occur for reasons discussed in subsection 5.2.2 – simultaneous right- and left-adjunction to VP$_4$ is excluded, and if XP$_1$ is scrambled to IP before XP$_2$ is extraposed to VP$_4$ (as required by the Barriers Condition), the Strict Cycle Condition is violated.

The same point can be made with instances of long-distance remnant topicalization. The examples in (133) show that successive-cyclic long-distance remnant topicalization is an option in German (see subsection 3.3.3 of chapter 5 for more discussion):

(133) a. [$_{VP}$ t$_2$ Gerechnet]$_3$ denke ich [$_{CP}$ t$'_3$ dass keiner [$_{PP}$ da$_1$mit]$_2$ t$_3$ hat
counted think I that no-one therewith has

b. [$_{VP}$ t$_2$ Gelesen]$_3$ glaubt sie [$_{CP}$ t$'_3$ dass keiner [$_{NP}$ ein Buch darüber$_1$]$_2$
read thinks she that no-one a book about that
t$_3$ hat
has

Here, the embedded clause is verb-final; the auxiliary *hat* stays in I. Given the presence of the auxiliary in I, XP$_2$ cannot possibly be right-adjoined to CP in these sentences; rather, XP$_2$ must have undergone either string-vacuous scrambling or local extraposition to a VP-adjoined position. This state of affairs predicts that PP$_2$ and NP$_2$ are islands for extraction in this construction; and as shown by the data in (134), this is indeed (by and large) the case:

(134) a. ??[$_{VP}$ t_2 Gerechnet $]_3$ denke ich [$_{CP}$ t'_3 dass da_1 keiner [$_{PP}$ t_1 mit $]_2$ hat
counted think I that there no-one with has

b. ?*[$_{VP}$ t_2 Gelesen $]_3$ glaubt sie [$_{CP}$ t'_3 dass $darüber_1$ keiner [$_{NP}$ ein Buch t_1 $]_2$ hat
read thinks she that about that no-one a book has

If the presence of a verb-final head (that precludes the option of analyzing the sentence via extraposition to CP) is responsible for the illformedness of the examples in (134), it should be the case that V/2 movement in the embedded clauses turns these examples into legitimate ones. As shown by the data in (135), this is so:

(135) a. [$_{VP}$ t_2 Gerechnet $]_3$ denke ich [$_{CP}$ t'_3 hat da_1 keiner [$_{PP}$ t_1 mit $]_2$
counted think I has there no-one with

b. [$_{VP}$ t_2 Gelesen $]_3$ glaubt sie [$_{CP}$ t'_3 hat $darüber_1$ keiner [$_{NP}$ ein Buch t_1 $]_2$
read thinks she has about that no-one a book

Thus, by and large, it seems to me that by reanalyzing string-vacuous scrambling as extraposition to CP in examples like those in (124), the Anti-Freezing effect with movement of XP_2 can be explained. This analysis raises a number of consequences, though, one of which I will now address.

Recall from subsection 7.2 that apparent cases of string-vacuous scrambling in remnant movement configurations do not exhibit categorial selectivity as normally occurs with scrambling; cf. the examples in (127). If string-vacuous scrambling can be reanalyzed as extraposition in these cases, the fact that restrictions on scrambling are apparently violated by the construction does not come as a surprise. However, it is also not the case that extraposition can affect any kind of category equally well. Thus, as noted in subsection 5.1.2, NPs can only undergo regular instances of extraposition in German if they are heavy (or focussed). Similarly, it turns out that secondary predicates and indefinite pronouns that take the form of *wh*-pronouns normally cannot undergo extraposition very well either. Thus, consider the contrast between the pair of sentences in (136), where regular extraposition is involved, and the pair of sentences in (137) (= (127)), where I have suggested that extraposition also occurs:

(136) a. ??dass der Karl das Fleisch t_1 gegessen hat [$_{AP}$ roh $]_1$
that ART Karl the meat eaten has raw

b. ?*dass gestern Antje t_1 gelesen hat was_1
that yesterday Antje read has something

(137) a. [$_{VP}$ t_1 Gegessen $]_2$ hat der Karl das Fleisch t_2 [$_{AP}$ roh $]_1$
eaten has ART $Karl_{nom}$ the $meat_{acc}$ raw

b. [$_{VP}$ t_1 Gelesen]$_2$ hat Antje gestern t_2 was$_1$
read has Antje yesterday something

Under present assumptions, this means that the (im-) possibility of extraposition of some item α cannot be an inherent property of α, but must be related to the syntactic context. Indeed, it seems that recent research according to which extraposition options are at least partly determined by phonological conditions (cf. in particular Truckenbrodt (1994), but also Büring & Hartmann (1997); also see the problem noted in footnote 22 above) might offer a clue as to why extraposition is severely degraded in (136), but possible in (137). I will not pursue the matter here, but it is clear that some account along the lines just indicated must be given for the contrast between (136) and (137) if the approach to unexpected Anti-Freezing with (apparent) string-vacuous scrambling in remnant movement configurations is to be successful.[53]

As before, many questions are left open; in particular, it seems that the approach to (Anti-) Freezing and remnant movement developed here is still somewhat too permissive (which is ultimately due to the fact that the proposed system permits ample left- and right-adjunction).[54] However, I will leave it at that, again emphasizing the

[53]Admittedly, this reasoning implies that the lack of categorial selectivity noted in subsection 7.2 could in principle also be reconciled with the idea that what is involved here is always string-vacuous scrambling. Under this view, however, one would have to motivate the idea that phonological factors are responsible for the weak ban on, e.g., predicate scrambling in normal contexts – a task that strikes me as harder than the one just envisaged.

[54]Let me just mention one construction that is no longer blocked under the assumptions made in this section. Consider the sentences in (i):

(i) a. ?*[$_{VP}$ [$_{VP}$ Da$_1$ [$_{VP}$ t_2 gerechnet]$_3$]$_3$ [$_{PP}$ t_1 mit]$_2$]$_3$ hat der Fritz nicht t_3
there counted with has ART Fritz not

b. *[$_{VP}$ Da$_1$ [$_{VP}$ t_2 gerechnet]$_3$]$_3$ hat der Fritz [$_{PP}$ t_1 mit]$_2$ nicht t_3
there counted has ART Fritz with not

c. ?*[$_{VP}$ Da$_1$ [$_{VP}$ t_2 gerechnet]$_3$]$_3$ hat der Fritz nicht [$_{PP}$ t_1 mit]$_2$
there counted has ART Fritz not with

Recall that a combination of PP extraposition and NP scrambling to the same VP, as in (i-a) (= (96)), is ruled out with the help of assumption (94). (i-b) is also excluded. Here, a scrambled PP contains a trace of an NP that has also undergone scrambling, which is impossible (see the following chapter for extensive discussion). However, (i-c) (which is taken from den Besten & Webelhuth (1990, 82)) now poses a problem, given that right-peripheral categories like PP_2 in (i-c) do not have to be analyzed in terms of string-vacuous scrambling (in which case the ungrammaticality would follow on a par with (i-b)), but can show up in their surface positions as a result of extraposition to CP. Thus, as it stands, (i-c) can have a derivation that does not violate any constraints: Scrambling of NP_1 to VP is followed by right-adjunction of PP_2 to CP (TopP), which in turn is followed by remnant VP topicalization. To solve this problem, additional assumptions appear necessary. Still, it should be noted that analogous constructions involving not NP scrambling from PP, but PP scrambling from NP are not as deviant; thus, the contrast that fails to show up (in a strong form, at least) in (i-bc) does occur in (ii-ab):

(ii) a. *[$_{VP}$ Darüber$_1$ [$_{VP}$ t_2 gelesen]$_3$]$_3$ hat [$_{NP}$ ein Buch t_2] keiner t_3
about that read has a book no-one

speculative nature of this section, and come to a conclusion.

8. Conclusion

Let me summarize the main findings of this chapter. First, and perhaps most importantly, the systematic Anti-Freezing effect that arises with typical instances of remnant movement is accounted for in a straightforward way without invoking construction-specific assumptions (e.g., about unbound traces). Given the derivational approach to movement developed in chapter 2, the non-occurrence of Freezing with remnant movement is exactly what we are led to expect. Second, it has turned out that the Anti-Freezing effect that regularly arises with extraposition (of CP, PP, and NP) in SOV languages is amenable to the same kind of analysis, given the hypothesis that right-adjunction is possible to any kind of XP, including CP. According to this view, extraposition is possible without inducing a Freezing effect because it can be remnant movement. Third, I have advanced some pieces of independent corroboration of the analysis. One particularly interesting prediction of the present approach is that extraposition is expected to create a Freezing effect if it is not remnant movement, i.e., if it adjoins in a lower position than the preceding extraction operation (e.g., to VP). This prediction was shown to be borne out. Fourth, it has turned out that the account of Freezing effects with "low" extraposition requires a modification of the notion of minimal residue, to the effect that only the positions X-Adj, XP-Adj, and SpecX are in the minimal residue of X (i.e., that there is no recursion); this modification was shown to be independently called for. Fifth and finally, I have argued that data involving successive-cyclic *wh*-movement from multiply embedded CPs in German show that chain interleaving is not per se excluded (as predicted under the standard notion of Fewest Steps); rather, it is excluded only if it would induce improper movement, i.e., a PUB effect (as predicted under the revised versions of Fewest Steps and Last Resort argued for in chapter 2, for independent reasons). One thing that I left open in this chapter is how the apparent ban on (many cases of) remnant scrambling is to be derived. I turn to this issue and to other movement type asymmetries that show up with remnant movement in the following chapter.

---

b. ?[$_{VP}$ Darüber$_1$ [$_{VP}$ t$_2$ gelesen]$_3$]$_3$ hat keiner t$_3$ [$_{NP}$ ein Buch t$_2$]
about that read has no-one a book

The contrast in (ii) is what we are led to expect under present assumptions (since NP$_2$ can be extraposed in (ii-b), but not in (ii-a)). Judgements on all these data are fairly subtle, though, and vary a lot from speaker to speaker, so that it is unclear what theoretical conclusions are to be drawn from this evidence.

Chapter 5

Remnant Movement Asymmetries

1. Introduction

This chapter addresses the third peculiar property of remnant movement that was noted in section 3 of chapter 1, viz., that not all kinds of movement may affect remnant categories in the same way. More specifically, upon first glance it looks as though whereas remnant topicalization and remnant *wh*-movement are possible, remnant scrambling and remnant left dislocation are not. Since all these movement types may perfectly well affect complete, i.e., non-remnant categories, this might be taken to indicate that the behaviour of remnant categories differs significantly from that of other categories and that construction-specific constraints blocking illicit cases of remnant movement are called for. However, I will argue that this is not the case: First, the movement type asymmetries with remnant movement can be shown to be only apparent – upon closer scrutiny, it emerges that any movement type can in principle affect remnant XPs and that any movement type can be blocked under certain circumstances (having to do with the position of the antecedent of the unbound trace). And second, the observed constraint on remnant movement can be made to follow from a general principle of Unambiguous Domination that covers bound and unbound traces alike (i.e., that does not apply specifically to remnant movement constructions) and that can be shown to be derivable, at least to a large extent, from a modification of Last Resort that is independently motivated because it integrates the residue of the Shortest Path Condition (viz., the Minimal Link Condition of Chomsky (1995)).

I proceed as follows: In section 2, the apparent remnant movement asymmetries (topicalization and *wh*-movement vs. scrambling and left dislocation) are laid out. Section 3 presents conflicting evidence in the guise of illicit examples of remnant *wh*-movement and legitimate instances of remnant scrambling and remnant left dislocation. Section 4 introduces the principle of Unambiguous Domination that accounts for the illegitimate cases of remnant movement discussed so far. Sections 5 and 6 contain some refinements of the proposed analysis, and section 7 finally shows how the main effects of Unambiguous Domination can be derived from a revised version of

Last Resort.

2. The Phenomenon

2.1. Remnant Topicalization vs. Remnant Scrambling

From the discussion in the preceding chapters, it is clear that topicalization can affect remnant infinitives (coherent infinitives), remnant VPs, remnant NPs, and remnant APs. However, if one looks at minimally different examples involving scrambling of these remnant XPs, the result is ill formed.

2.1.1. Coherent Infinitives

This asymmetry between topicalization and scrambling has been observed for coherent infinitives (i.e., remnant infinitives from which scrambling has taken place) by a number of people, among them Fanselow (1991, 133; 1992), Grewendorf & Sabel (1994, 284ff), Frank, Lee & Rambow (1992, 152), Stechow (1992, 222ff), and Haider (1993, 234ff). Relevant examples that show this apparent contrast between topicalization and scrambling are given in (1) and (2):[1]

(1) a. [$_\alpha$ t_1 Zu lesen]$_3$ hat [$_{NP}$ das Buch]$_1$ keiner t_3 versucht
to read has the book no-one tried

b. [$_\alpha$ t_1 t_2 Zu reparieren]$_3$ hat der Frank dem Matthias$_1$ den Drucker$_2$ t_3 versprochen
to fix has ART Frank ART Matthias$_{dat}$ the printer$_{acc}$ promised

(2) a. *dass [$_\alpha$ t_1 zu lesen]$_3$ [$_{NP}$ das Buch]$_1$ keiner t_3 versucht hat
that to read the book$_{acc}$ no-one tried has

b. *dass der Frank [$_\alpha$ t_1 t_2 zu reparieren]$_3$ dem Matthias$_1$ den Drucker$_2$ t_3 versprochen hat
that ART Frank to fix ART Matthias$_{dat}$ the printer$_{acc}$ promised has

As noted in subsection 3.3 of chapter 1, this asymmetry breaks down in the case of complete infinitives. Both topicalization and scrambling can affect the infinitive if it does not contain a trace that becomes unbound after movement; cf. (3) and (4).

(3) a. [$_\alpha$ Das Buch zu lesen]$_3$ hat keiner t_3 versucht
the book$_{acc}$ to read has no-one tried

---

[1] As before, I remain uncommitted as to the identity of α; see chapter 1.

b. [$_\alpha$ Dem Matthias den Drucker zu reparieren]$_3$ hat der Frank t$_3$
ART Matthias$_{dat}$ the printer$_{acc}$ to fix has ART Frank$_{nom}$
versprochen
promised

(4) a. dass [$_\alpha$ das Buch zu lesen]$_3$ keiner t$_3$ versucht hat
that the book to read no-one tried has

b. dass [$_\alpha$ dem Matthias den Drucker zu reparieren]$_3$ der Frank t$_3$
that ART Matthias$_{dat}$ the printer$_{acc}$ to fix ART Frank
versprochen hat
promised has

2.1.2. Remnant VPs

Now consider remnant VPs. Remnant VP topicalization is, of course, the construction that Thiersch (1985) and den Besten & Webelhuth (1987) motivated remnant movement for in the first place (cf. chapter 1). Some examples are given in (5):

(5) a. [$_{\text{VP}}$ t$_1$ Gelesen]$_3$ hat das Buch$_1$ keiner t$_3$
read has the book no-one

b. [$_{\text{VP}}$ Dem Peter t$_2$ gegeben]$_3$ hat die Claudia ein Buch$_2$ t$_3$
ART Peter$_{dat}$ given has ART Claudia$_{nom}$ a book$_{acc}$

c. [$_{\text{VP}}$ t$_1$ t$_2$ Gegeben]$_3$ hat die Claudia dem Peter$_1$ ein Buch$_2$ t$_3$
given has ART Claudia$_{nom}$ ART Peter$_{dat}$ a book$_{acc}$

d. [$_{\text{VP}}$ Dem Peter ein Buch gegeben]$_3$ hat die Claudia gestern t$_3$
ART Peter$_{dat}$ a book$_{acc}$ given has ART Claudia yesterday

In (5-a) (cf. (1-a) from chapter 1), a VP that contains the trace of a scrambled direct object is topicalized. (5-b) shows that in VPs headed by triadic verbs, a direct object trace may occur unbound after topicalization. Furthermore, it can even be the case that both arguments in a double object construction undergo scrambling prior to VP-topicalization, as in (5-c). Finally, (5-d) instantiates a case of topicalization of a complete, i.e., non-remnant, VP with two objects.

In contrast, scrambling of remnant VPs results in ungrammaticality; this is shown in (6-a) through (6-c) (also see Grewendorf (1995, 1306-1307)).

(6) a. *dass [$_{\text{VP}}$ t$_1$ gelesen]$_3$ das Buch$_1$ keiner t$_3$ hat
that read the book no-one has

b. *dass [$_{\text{VP}}$ dem Peter t$_2$ gegeben]$_3$ die Claudia ein Buch$_2$ t$_3$ hat
that ART Peter$_{dat}$ given ART Claudia a book$_{acc}$ has

c. *dass [$_{\text{VP}}$ t$_1$ t$_2$ gegeben]$_3$ die Claudia dem Peter$_1$ ein Buch$_2$ t$_3$ hat
that given ART Claudia ART Peter$_{dat}$ a book$_{acc}$ has

The strong illformedness of the examples in (6) is evidently due to the fact that an unbound trace occurs in the scrambled VP. If a complete VP that does not contain an object trace undergoes scrambling, as in (7), the resulting sentence improves significantly:

(7) ??dass [$_{VP}$ dem Peter ein Buch gegeben]$_3$ die Claudia gestern nicht t_3
that ART Peter$_{dat}$ a book$_{acc}$ given ART Claudia yesterday not
hat
has

Still, as indicated, (7) is far from being perfect. Rather (as was noted in footnote 14 of chapter 4), it has an intermediate status and is possible only if a certain intonation pattern is present, viz., what Jacobs (1982) calls *I-topicalization* ("intonational topicalization"). I-topicalization is typically characterized by a rising tone on the moved item (in this case, on the direct object immediately preceding the verb of the moved VP) and a falling tone on a constituent that follows the moved item but generally precedes the trace of the moved item (here, on the negation *nicht* ('not')).[2] At present it is immaterial why exactly VP scrambling (more generally, predicate scrambling; see the following subsection) is per se somewhat marginal in German and why it requires the I-topicalization intonation pattern to be acceptable at all (see Webelhuth (1987), Stechow & Sternefeld (1988, 464ff), Reis & Rosengren (1992), Haider (1993, 197ff), and Müller (1995, 153-158)). What is relevant, though, is that despite the intermediate status that is generally associated with all kinds of predicate scrambling in German, there is still a clear contrast to cases of remnant predicate scrambling as in (6); it is this contrast that is in need of an explanation.

2.1.3. Remnant APs

Let me now turn to APs. Note first that remnant APs can undergo topicalization in German:

(8) [$_{AP}$ Stolz t_1]$_2$ ist der Fritz gestern [$_{PP}$ auf sein Kind]$_1$ nicht t_2 gewesen
proud is ART Fritz yesterday of his child not been

Again, however, remnant AP movement is impossible if the movement type affecting AP is scrambling. This is shown by the illformedness of (9-a) and (9-b):

(9) a. *dass [$_{AP}$ stolz t_1]$_2$ der Fritz [$_{PP}$ auf sein Kind]$_1$ nicht t_2 gewesen ist
that proud ART Fritz of his child not been is
b. *dass der Fritz [$_{AP}$ stolz t_1]$_2$ nicht [$_{PP}$ auf sein Kind]$_1$ t_2 gewesen ist
that ART Fritz proud not of his child been is

As was the case with VPs, AP scrambling in German has an intermediate status and requires the presence of an I-topicalization pattern to be acceptable at all (this shows

[2]Often, I-topicalization indicates reconstruction or "scope inversion," but this does not always seem to be the case. See Büring (1996) for extensive discussion.

that the restriction in question is one on predicates in general, and not simply on VPs); but the examples in (10), which involve scrambling of a complete AP, are much better than those in (9), which again strongly suggests that there is a specific constraint on remnant movement at work in (9) that overrides the marginality ascribable to predicate scrambling as such.[3]

(10) a. ?dass $[_{AP}$ stolz auf sein Kind $]_2$ der Fritz nicht t_2 gewesen ist
that proud of his child ART Fritz not been is

b. ??dass der Fritz $[_{AP}$ stolz auf sein Kind $]_2$ nicht t_2 gewesen ist
that ART Fritz proud of his child not been is

2.1.4. Remnant NPs

Finally, recall that topicalization of a remnant NP is possible in German:

(11) a. $[_{NP}$ Ein Buch t_1 $]_2$ hat Antje $[_{PP}$ über die Liebe $]_1$ t_2 gelesen
a book has Antje about the love read

b. $[_{NP}$ Ein Buch t_1 $]_2$ hat $[_{PP}$ über die Liebe $]_1$ niemand t_2 gelesen
a book has about the love no-one read

In both examples, a PP argument is extracted from an object NP. In (11-a) the PP has undergone scrambling to VP; in (11-b) it is adjoined to IP. In addition, the remnant NP is topicalized in both cases without inducing ungrammaticality.

In contrast to that, scrambling of a remnant NP invariably results in ungrammaticality. Consider (12):

(12) *dass niemand $[_{NP}$ ein Buch t_1 $]_2$ gestern $[_{PP}$ über die Liebe $]_1$ t_2 gelesen hat
that no-one a book yesterday about the love read has

Here, a remnant NP from which a PP has been extracted and adjoined to VP via scrambling is also adjoined to VP, producing illformedness. Similarly, IP-adjunction of a remnant NP is prohibited:

(13) *dass $[_{NP}$ ein Buch t_1 $]_2$ niemand $[_{PP}$ über die Liebe $]_1$ t_2 gelesen hat
that a book no-one about the love read has

The ungrammaticality of (13) cannot be due to the fact that an indefinite NP is scrambled in front of the subject – bare negative quantifiers in SpecI, unlike some other kinds of NPs in German, tolerate scrambling of an indefinite NP in front of the subject position. Thus, if an NP that does not contain an unbound trace is scrambled in the

[3]Notice also that I-topicalization does not help to improve the examples in (9) and (6).

same context, the result is fine:[4]

(14) dass [$_{NP}$ ein Buch [$_{PP}$ über die Liebe]]$_2$ niemand t$_2$ gelesen hat
that a book about the love no-one read has

Summarizing so far, a very simple generalization seems to emerge – remnant categories can be topicalized, but they cannot undergo scrambling.[5] In the following two subsections, I consider two further movement types, viz., *wh*-movement and left dislocation, in their behaviour with respect to remnants.

2.2. Remnant Wh-Movement

As shown by the examples in (15), remnant NPs can undergo *wh*-movement to SpecC in German:[6]

(15) a. [$_{NP}$ Was für ein Buch t$_1$]$_2$ hast du [$_{PP}$ über die Liebe]$_1$ t$_2$ gelesen ?
what for a book have you about the love read

---

[4]This implies that the weak ban on predicate scrambling that creates an intermediate status in all cases of VP and AP scrambling does not extend to NPs, just as it did not affect coherent infinitives – the latter two categories function as arguments in the examples under discussion, not as predicates.

[5]The prohibition against remnant scrambling is also active in languages like Japanese (cf. Saito (1992, 83), Kitahara (1994, 203-205), and Koizumi (1995, 185-194), among others) and Korean (Shin-Sook Kim (p.c.)). For reasons which will not concern me here (but cf. Müller & Sternefeld (1993; 1996)), Japanese and Korean, in contrast to German (see the discussion in subsection 4.2 of chapter 2), permit long-distance scrambling of arguments from finite clauses, as in the Japanese example (i):

(i) [$_{NP}$ Sono hon-o]$_1$ John-ga [$_{CP}$ Mary-ga t$_1$ yonda-to]$_2$ itta (koto)
that book$_{acc}$ John$_{nom}$ Mary$_{nom}$ read-COMP said fact

It is also possible that a whole embedded CP undergoes long-distance scrambling in front of the matrix subject, as in the Japanese example (ii-a). However, what may never occur is a combination of long-distance scrambling of an NP from an embedded CP, followed by scrambling of the remnant CP in front of this NP, as in (ii-b) (both examples are taken from Kitahara (1994, 203)):

(ii) a. [$_{CP}$ Mary-ga [$_{NP}$ sono hon-o]$_1$ yonda-to]$_2$ Bill-ga [$_{CP}$ John-ga t$_2$ itta-to]
Mary$_{nom}$ that book$_{acc}$ read-COMP Bill$_{nom}$ John$_{nom}$ said-COMP
omotteiru (koto)
think fact

b. *[$_{CP}$ Mary-ga t$_1$ yonda-to]$_2$ [$_{NP}$ sono hon-o]$_1$ John-ga t$_2$ itta (koto)
Mary$_{nom}$ read-COMP that book$_{acc}$ John$_{nom}$ said fact

Thus, however one derives the illicit cases of remnant scrambling in German, it seems reasonable to postulate that cases like (ii-b) in Japanese (and Korean) should be excluded in the same way.

[6]Recall that these examples have already been briefly discussed in subsection 3.3 of chapter 1 and then, in somewhat more detail, in subsection 4.2.2 of chapter 3 (with respect to the Proper Binding Condition) and in subsection 3.1 of chapter 4 (with respect to Freezing). Recall also that the slight deviance of (15-b) was attributed in chapter 3 to a Specificity effect induced by the scrambling operation that removes the PP *über die Liebe* ('about love') from the definite *which*-NP.

b. ?[$_{NP}$ Welches Buch t_1]$_2$ hat [$_{PP}$ über die Liebe]$_1$ niemand t_2 gelesen ?
which book has about the love no-one read

Similarly, remnant APs can undergo *wh*-movement, just as they can undergo topicalization. This is shown in (16):

(16) [$_{AP}$ Wie stolz t_1]$_2$ ist der Fritz [$_{PP}$ auf sein Kind]$_1$ gestern t_2 gewesen ?
how proud is ART Fritz of his child yesterday been

Thus, the conclusion can be drawn that remnant *wh*-movement in German patterns with remnant topicalization and not with remnant scrambling.

Given this state of affairs, an interesting prediction arises for cases of apparent clausal pied piping of infinitives with relativization in German, as in (17-a).

(17) a. ein Freund [$_{CP_4}$ dem$_1$ t_1 das Fahrrad zu reparieren der Fritz
a friend who$_{dat}$ the bicycle$_{acc}$ to fix ART Fritz$_{nom}$
versprochen hat]
promised has

b. ... [$_{CP_4}$ [$_{\alpha_2}$ dem$_1$ t_1 das Fahrrad zu reparieren] C der Fritz t_2 versprochen hat]

c. ... [$_{CP_4}$ dem$_1$ C [$_{\alpha_2}$ t_1 das Fahrrad zu reparieren] der Fritz t_2 versprochen hat]

There is disagreement as to whether (17-a) should be analyzed in terms of genuine pied piping of the infinitive, i.e., as movement of a complete infinitive to the specifier of the relative clause CP, as in (17-b) (cf. van Riemsdijk (1985)), or in terms of a co-occurrence of scrambling of the infinitive to IP and movement of the relative pronoun to the specifier of the relative clause CP_4, as in (17-c) (cf. Grewendorf (1988)). With this in mind, consider (18-a):

(18) a. *ein Freund [$_{CP}$ dem$_1$ t_1 t_3 zu reparieren das Fahrrad$_3$ der Fritz
a friend who$_{dat}$ to fix the bicycle$_{acc}$ ART Fritz$_{nom}$
versprochen hat]
promised has

b. [$_\alpha$ t_1 t_2 Zu reparieren] hat der Fritz dem Frank$_1$ das Fahrrad$_2$
to fix has ART Fritz$_{nom}$ ART Frank$_{dat}$ the bicycle$_{acc}$
versprochen
promised

c. *dass [$_\alpha$ t_1 t_2 zu reparieren] der Fritz dem Frank$_1$ das Fahrrad$_2$
that to fix ART Fritz$_{nom}$ ART Frank$_{dat}$ the bicycle$_{acc}$
vesprochen hat
promised has

In (18-a), fronting of an incomplete infinitive takes place in the pied piping construction (i.e., of an infinitive that contains an unbound scrambling trace), and the result is ungrammatical (cf. van Riemsdijk (1985, 182) for the original observation). This is to be expected if cases of apparent clausal pied piping in German involve scrambling, as proposed by Grewendorf (1988) (cf. (18-c)), but necessitates additional assumptions if real pied piping is involved, as suggested by van Riemsdijk (1985) – in the latter approach, remnant infinitive fronting in (18-a) is viewed as a *wh*-movement (-like) operation, and it should therefore behave like other instances of specifier movement (such as topicalization, cf. (18-b)), rather than like scrambling.

2.3. Remnant Left Dislocation

Truckenbrodt (1992), elaborating on some pertinent remarks in Haider (1990, 100f), observes that whereas topicalization of remnant VPs is possible, left dislocation of remnant VPs is not. The illegitimacy of remnant VP left dislocation in German is shown in (19):[7]

(19) a. *[$_{\text{VP}}$ t_1 Gelesen]$_{,}$ das hat das Buch_1 keiner t
read that has the book no-one

b. *[$_{\text{VP}}$ Dem Peter t_2 gegeben]$_{,}$ das hat die Claudia ein Buch_2 t
ART Peter_{dat} given that has ART Claudia_{nom} a book_{acc}

c. *[$_{\text{VP}}$ t_1 t_2 Gegeben]$_{,}$ das hat die Claudia dem Peter_1 ein Buch_2 t
given that has ART Claudia_{nom} ART Peter_{dat} a book_{acc}

In these examples, there is left dislocation of remnant VPs, resulting in ungrammaticality; (20), on the other hand, shows that left dislocation of a complete VP (including two objects in this case) is possible:

(20) [$_{\text{VP}}$ Dem Peter ein Buch gegeben]$_{,}$ das hat die Claudia gestern t
ART Peter_{dat} a book_{acc} given that has ART Claudia yesterday

---

[7] Note that the left dislocation data in (19) differ from their topicalization counterparts in (5) only in that a resumptive pronoun *das* ('that') is present in the former and absent in the latter. It seems to me that the asymmetry between topicalization and left dislocation that can be observed here casts serious doubts on attempts to reduce one movement type to the other in German, as is suggested, inter alia, by Koster (1978), Cardinalletti (1986), Zwart (1993), and van Hoof (1995) (partially for Dutch) – topicalization and left dislocation behave differently in German. That said, it should be noted that a strong asymmetry between topicalization and left dislocation with respect to remnant movement does not seem to show up in Dutch; consider, for instance, the following examples from Zwart (1993, 109):

(i) a. [$_{\text{VP}}$ t_1 Gekust]$_2$ heeft Jan Marie_1 niet t_2
kissed has John_{nom} Marie_{acc} not

b. [$_{\text{VP}}$ t_1 Gekust]$_2$ dat heeft Jan Marie_1 niet t_2
kissed that has John_{nom} Marie_{acc} not

For the time being, I will abstract away from these data; but see footnote 32.

This pattern is very similar to that found with VP scrambling (abstracting away from the intervening predicate scrambling effect).[8]

Indeed, left dislocation seems to behave exactly like scrambling in this respect. Thus, left dislocation of a remnant NP (as in (21-a), cf. the corresponding topicalization case in (11-a)), a remnant AP (as in (21-b), cf. topicalization in (8)), or a coherent infinitive (as in (21-c), cf. topicalization in (1-a)) is also impossible:

(21) a. *[$_{\text{NP}}$ Ein Buch t$_1$]$_,$ das hat Antje [$_{\text{PP}}$ über die Liebe]$_1$ t gelesen
a book that has Antje about the love read

b. ?*[$_{\text{AP}}$ Stolz t$_1$]$_,$ das ist der Fritz gestern [$_{\text{PP}}$ auf sein Kind]$_1$ nicht t gewesen
proud that is ART Fritz yesterday of his child not been

c. *[$_\alpha$ t$_1$ Zu lesen]$_,$ das hat [das Buch]$_1$ keiner t versucht
to read that has the book no-one tried

As shown in (22), minimally different examples in which a complete NP, AP, or infinitive undergoes left dislocation are perfectly grammatical:

(22) a. [$_{\text{NP}}$ Ein Buch über die Liebe]$_,$ das hat Antje t gelesen
a book about the love that has Antje read

---

[8] In (5), (6) and (19), I have systematically ignored examples involving movement of a VP that contains only the verb and a direct object in double object constructions. Indeed, in this case, remnant scrambling is somewhat better than in (6-a) – (6-c) and is more on a par with (7). Similarly, as noted by Truckenbrodt (1992), remnant left dislocation of VPs containing only a verb and a direct object is fully well formed, and thus patterns with (20) rather than with the sentences in (19). Cf. topicalization in (i-a), scrambling in (i-b), and left dislocation in (i-c):

(i) a. [$_{\text{VP}}$ (t$'_1$) Ein Buch (t$_1$) gegeben]$_3$ hat die Claudia dem Peter$_1$ t$_3$
a book$_{acc}$ given has ART Claudia ART Peter$_{dat}$

b. ??dass [$_{\text{VP}}$ (t$'_1$) ein Buch (t$_1$) gegeben]$_3$ die Claudia dem Peter$_1$ nicht t$_3$ hat
that a book$_{acc}$ given ART Claudia ART Peter$_{dat}$ not has

c. [$_{\text{VP}}$ (t$'_1$) Ein Buch (t$_1$) gegeben]$_,$ das hat die Claudia dem Peter$_1$ nicht t
a book$_{acc}$ given that has ART Claudia ART Peter$_{dat}$ not

Taken at face value, this might be regarded as an indication that the VPs in (i-abc) do not contain a trace of the indirect object. On the other hand, if one assumes (following Larson (1988)) that dative-bearing indirect objects have undergone movement to a Case-position from a position that is asymmetrically c-commanded by the direct object position at D-structure, then the moved VPs in (i) must be genuine remnant XPs containing unbound traces, at least in the base position c-commanded by the direct object and possibly even in the position to which the indirect object is dative-moved in order to receive Case (if this position is VP-internal, as suggested by Larson). Under these assumptions, a problem would arise for the generalization that remnants cannot undergo scrambling or left dislocation in German. For the time being, I ignore these potential complications; I will return to examples like (i-b) and (i-c) in section 5 below and show that they do not pose a problem for the approach I am going to advance.

b. [$_{AP}$ Stolz auf sein Kind]$_i$ das ist der Fritz gestern nicht t gewesen
   proud of his child that is ART Fritz yesterday not been

c. [$_\alpha$ Das Buch zu lesen]$_i$ das hat keiner t versucht
   the book to read that has no-one tried

Thus, remnant left dislocation in German patterns with remnant scrambling, and not with remnant topicalization and remnant *wh*-movement.

2.4. A Preliminary Generalization

To sum up so far, the following preliminary generalization suggests itself:

(23) *Generalization* (preliminary):
Remnant XPs can undergo topicalization and *wh*-movement, but not scrambling or left dislocation.

According to (23), it is an inherent property of a movement type to either license or prohibit remnants. And indeed, attempts have been made to derive such apparent movement type asymmetries as they are documented in (23). Thus, e.g., Frank, Lee & Rambow (1992), Stechow (1992), and Grewendorf & Sabel (1994), all noting the contrast between topicalization and scrambling in the case of coherent infinitives, try to identify intrinsic properties of these two movement types that lead to wellformedness in the case of remnant topicalization and to illformedness in the case of remnant scrambling.[9] Along the same lines, Truckenbrodt (1992), observing the asymmetry between remnant VP topicalization and remnant VP left dislocation, assumes that it is an inherent property of left dislocation not to permit remnant movement, whereas no such ban is active with topicalization.[10]

However, notwithstanding some theory-internal problems with all these approaches that I will not go into here (cf. Müller (1993, 413-422)), I believe that the generali-

[9]Both Stechow (1992) and Frank, Lee & Rambow (1992) suggest that the relevant factor is the Proper Binding Condition (see chapter 3) and that unbound traces in topicalized items can fulfill this condition (either via LF reconstruction, as in Stechow's approach, or via chain-binding, as in Frank, Lee & Rambow's approach), whereas unbound traces in scrambled items cannot. In contrast, Grewendorf & Sabel (1994) propose that the relevant factor is the notion of barrier, and they go on to develop a theory of locality according to which scrambled items are invariably barriers, whereas topicalized items may or may not be barriers, depending on their position, the lexical properties of the embedding predicate, etc.

[10]The actual implementation of this approach relies again on the Proper Binding Condition, assumed to hold at LF; Truckenbrodt (1992) suggests that topicalized items are always reconstructed, so that unbound traces can be prevented at LF (cf. subsection 4.1 of chapter 3), whereas left dislocated items are never reconstructed, so that unbound traces created by remnant left dislocation remain at LF, in violation of the proper binding requirement.

zations they are based on (viz., some partial statement of (23)) are indeed untenable.[11] The reason is that closer inspection reveals that the situation is more complicated: The ability or inability to affect remnant categories is in fact not an intrinsic property of a given movement type. A first piece of evidence that points in this direction arises under the Predicate-Internal Subject Hypothesis.

---

[11] As noted, several approaches to the apparent movement type asymmetries rely on a version of the Proper Binding Condition (PBC). In chapter 3, I have argued that all versions of the PBC should be dispensed with, leaving only a semantically motivated residue of this condition, viz., the requirement that semantic variables are syntactically bound at LF (see subsection 4.3 of chapter 3). But irrespective of this, let me mention just one observation that casts doubt on any approach to the apparent movement type asymmetries that relies on something like the PBC. This observation is related to the behaviour of seemingly unbound anaphoric items, like the reflexive pronoun *sich*, in moved XPs (see especially Barss (1984; 1986) and Saito (1989) for relevant discussion in this context). As shown in (i), *sich* can occur in a scrambled NP, AP, infinitive, or VP (note that (i-b) and (i-d) are degraded for independent reasons because they involve predicate scrambling):

(i) a. dass [$_{\text{NP}}$ ein Buch über sich$_1$]$_2$ niemand$_1$ t$_2$ gerne lesen würde
that a book about REFL no-one happily read would

b. ?dass [$_{\text{AP}}$ stolz auf sich$_1$]$_2$ der Fritz$_1$ nie t$_2$ gewesen ist
that proud of REFL ART Fritz never been is

c. dass [$_\alpha$ sich$_1$ zu waschen]$_2$ keiner$_1$ mehr t$_2$ versucht hat
that REFL to wash no-one anymore tried has

d. ??dass [$_{\text{VP}}$ sich$_1$ dem Peter gezeigt]$_2$ letzte Woche die Claudia$_1$ t$_2$ hat
that REFL ART Peter$_{dat}$ shown last week ART Claudia$_{nom}$ has

Thus, there is a striking contrast between unbound lexical anaphors and unbound traces in scrambled items. Assuming both anaphors and traces to be subject to the PBC, the contrast between the data in (i) and corresponding examples involving remnant scrambling is unexplained. (Note, however, that under the Predicate-Internal Subject Hypothesis, the seemingly unbound anaphor *sich* would be bound even at S-structure in (i-bcd); under this hypothesis, only the NP scrambling case in (i-a) would make the point. I return to this issue directly.)

Basically the same conclusion applies in the case of left dislocation – as with scrambling, the pertinent examples (cf. (19) & (21)) improve a great deal if an unbound anaphor replaces a trace in the XP which has undergone left dislocation:

(ii) a. [$_{\text{NP}}$ Ein Buch über sich$_1$], das hat der Fritz$_1$ t gelesen
a book about REFL that has ART Fritz read

b. [$_{\text{AP}}$ Stolz auf sich$_1$], das ist Antje$_1$ jetzt endlich t
proud of REFL that is Antje now finally

c. [$_\alpha$ Sich$_1$ zu waschen], das hat dann keiner$_1$ mehr t versucht
REFL to wash that has then no-one anymore tried

d. [$_{\text{VP}}$ Sich$_1$ dem Peter gezeigt], das hat letzte Woche die Claudia$_1$ t
REFL ART Peter$_{dat}$ shown that has last week ART Claudia

(Again, only (i-a) is relevant under the Predicate-Internal Subject Hypothesis.) Thus, quite independently of the objections raised against the PBC in chapter 3, the fact that unbound lexical anaphors behave differently than unbound traces with respect to the issue of being contained in scrambled or left dislocated items strongly suggests that something like the PBC cannot be the correct means of deriving movement type asymmetries in the case of remnants.

2.5. Movement Type Asymmetries and Predicate-Internal Subjects

Suppose that subjects are base-generated VP-internally, as suggested by Kitagawa (1986), Larson (1988), Sportiche (1988), Speas (1990), Koopman & Sportiche (1991), Chomsky (1995), and many others. Suppose furthermore that they are generated in the SpecV position. Under these two assumptions, it is evident that an example like (24) (= (12) from chapter 1) must contain an unbound subject trace – according to this view, *John* has undergone Case-driven movement to SpecI prior to VP topicalization (cf. Roberts (1990, 392) for relevant discussion).

(24) $[_{VP}$ t_1 Kicked the dog $]_2$ $John_1$ never has t_2

Thus, as noted by Huang (1993) and Bayer (1993), given the Predicate-Internal Subject Hypothesis, there is a complete analogy between topicalization of the passivized VP in (25) (= (16) from chapter 1) and topicalization of the active voice VP in (24)– in both cases, an unbound trace shows up in the topicalized VP; the antecedent of the trace has undergone Case-driven movement to SpecI in the previous step.

(25) $[_{VP}$ Criticized t_2 by his boss $]_1$ $John_2$ has never been t_1

With this in mind, consider again a typical contrast between scrambling of an incomplete, remnant VP and scrambling of a complete VP in German, as in (26-a) (based on (6-a)) vs. (26-b) (where t_1 is the object trace, and t_2 the subject trace):[12]

(26) a. *dass $[_{VP}$ t_2 t_1 gelesen $]_3$ das $Buch_1$ $keiner_2$ t_3 hat
that read the book no-one has

b. ??dass $[_{VP}$ t_2 das $Buch_1$ gelesen $]_3$ $keiner_2$ t_3 hat
that the book read no-one has

As indicated, under the Predicate-Internal Subject Hypothesis, both (26-a) and (26-b) contain an unbound subject trace at S-structure. In addition, (26-a) has an unbound object trace, in contrast to (26-b). Evidently, then, there can be no prohibition against remnant scrambling as such. If the trace that becomes unbound as a result of remnant scrambling is a subject trace created by Case-driven movement, remnant scrambling is possible in principle; if it is an object trace created by scrambling, remnant movement is impossible. Thus, the conclusion to be drawn from the data in (26) is that any approach that postulates that it is an intrinsic property of the movement type scrambling not to apply to remnant XPs is doomed to fail under the Predicate-Internal Subject Hypothesis.

The same is true for left dislocation, given the contrast between (27-a) (cf. (19-a)) and (27-b) (again, t_1 is the object trace, and t_2 is the subject trace):

[12] As before, predicate scrambling as in (26-b) has an intermediate status and requires the presence of an I-topicalization intonational pattern – in this case, the rising tone is normally on *Buch* (‘book’), and the falling tone on *keiner* (‘no-one’).

(27) a. *[$_{\text{VP}}$ t$_2$ t$_1$ Gelesen]$_{,}$ das hat das Buch$_1$ keiner$_2$ t
read that has the book no-one

b. [$_{\text{VP}}$ t$_2$ Das Buch gelesen]$_{,}$ das hat keiner$_2$ t
the book read that has no-one

To sum up so far, the idea that certain movement types (like scrambling and left dislocation) simply can never affect remnant XPs, as an intrinsic property, forces one to reject the Predicate-Internal Subject Hypothesis. This may or may not be a fatal result; however, I take it that it points to a more general problem that I will discuss in the following section. There, I show that not only subject traces (as postulated under the Predicate-Internal Subject Hypothesis) can show up unbound in scrambled and left dislocated remnants, but also object traces under certain conditions; moreover, it emerges that object traces sometimes cannot show up in *wh*-moved (and topicalized) remnants, also in contrast to what we expect if these movement types systematically permit the creation of unbound traces in remnant XPs.

3. Conflicting Evidence

3.1. Remnant Scrambling is Sometimes Possible

In subsection 2.1.1, I gave sentences that involved illicit scrambling of a coherent infinitive, i.e., an infinitive from which scrambling has applied; relevant examples are given again in (28) ((28-b) = (2-a)).

(28) a. *dass [$_\alpha$ t$_1$ zu lesen]$_3$ keiner [das Buch]$_1$ t$_3$ versucht hat
that to read no-one the book$_{acc}$ tried has

b. *dass [$_\alpha$ t$_1$ zu lesen]$_3$ [das Buch]$_1$ keiner t$_3$ versucht hat
that to read the book$_{acc}$ no-one tried has

Interestingly, and completely unexpectedly under the generalization in (23), cases of remnant infinitive scrambling become acceptable if the antecedent of the unbound trace is not a full NP that has undergone scrambling, but rather a weak pronoun (cf. (29-a)) or a pronominal clitic (cf. (29-b)) in a Wackernagel (-like), i.e., pre-subject, position:

(29) a. dass [$_\alpha$ t$_1$ zu lesen]$_3$ es$_1$ keiner t$_3$ versucht hat
that to read it$_{acc}$ no-one tried has

b. dass [$_\alpha$ t$_1$ zu lesen]$_3$ ’s$_1$ keiner t$_3$ versucht hat
that to read it$_{acc}$ no-one tried has

I take this as an indication that pronoun movement to a pre-SpecI position in German is not an instance of scrambling; rather, some other movement type seems to be involved. Indeed, it has sometimes been argued (see, e.g., Thiersch (1978), Cardinaletti & Roberts (1991), Schmidt (1992), Zwart (1993), and Cardinaletti & Starke (1995)) that pronoun movements of the kind shown in (29) do not qualify as instances of scrambling but involve either A-movement to the specifier of some functional projection or head movement to a functional head. In what follows, I will assume that this view is essentially correct.[13]

Thus, it seems that there is nothing wrong with scrambling of coherent infinitives per se. Rather, what appears to be the crucial factor is the relation between the position of the remnant XP and the position of the antecedent of the unbound trace: If both these items are in scrambling positions, ungrammaticality results (as in (28-a) and in (28-b)); if the remnant occupies a scrambling position and the antecedent of the unbound trace occupies some other position (e.g., a Wackernagel (-like) position), then the result is well formed, ceteris paribus (cf. (29)).

This view is reinforced by the contrast between remnant VP scrambling as in (30-a) (= (6-a) from subsection 2.1.2) on the one hand and remnant VP scrambling as in (30-b) and (30-c) on the other:

(30) a. *dass [$_{VP}$ t_1 gelesen $]_3$ das Buch$_1$ keiner t_3 hat
that read the book no-one has

b. ??dass [$_{VP}$ t_1 gelesen $]_3$ es$_1$ keiner t_3 hat
that read it no-one has

c. ??dass [$_{VP}$ t_1 gelesen $]_3$ ’s$_1$ keiner t_3 hat
that read it no-one has

In the ill-formed sentence (30-a), both the remnant VP and the antecedent of the unbound trace *das Buch* (‘the book’) have undergone scrambling. In (30-b) and (30-c), however, only the remnant VP has undergone scrambling – the antecedent of the unbound trace is a weak pronoun or a clitic and thus occupies a different kind of position, by assumption. Therefore, the intermediate status of (30-b) and (30-c) again suggests that there is no general prohibition against scrambling of remnant XPs.[14,15]

---

[13]What exactly this functional projection related to Wackernagel-effects with pronouns looks like is not immediately relevant to the issue at hand, and I do not have any specific proposal to offer.

[14]As before, I assume that the fact that (30-b) and (30-c) are not completely grammatical is due to the weak prohibition against predicate scrambling. Accordingly, I-topicalization is necessary in these examples.

[15]In principle, we would expect the same improvement to arise with those cases of remnant scrambling where the antecedent of the unbound trace is not scrambled but extraposed. Indeed, I think there is a clear contrast between, e.g., (28-a), where NP$_1$ is scrambled, and (i-a), where NP$_1$ has undergone Heavy NP Shift, i.e., extraposition:

(i) a. ??dass [$_{\alpha}$ t_1 zu lesen $]_3$ keiner t_3 versucht hat [$_{NP}$ das letzte Buch [$_{CP}$ das der Fritz darüber
that to read no-one tried has the last book that ART Fritz about that
verfaßt hat $]]_1$
written has

3.2. Remnant Left Dislocation is Sometimes Possible

A similar case can be made with left dislocation. Data involving illicit remnant VP left dislocation, such as (19-a), which is repeated here as (31-a), sharply contrast with examples like (31-b) and (31-c).

(31) a. *[$_{VP}$ t_1 Gelesen], das hat das Buch$_1$ keiner t
read that has the book no-one

b. (?)[VP t_1 Gelesen], das hat es$_1$ keiner t
read that has it no-one

c. [$_{VP}$ t_1 Gelesen], das hat's$_1$ keiner t
read that has-it no-one

Again, the only relevant difference between the well-formed examples of remnant VP left dislocation in (31-b) and (31-c) and the ill-formed example in (31-a) is that the direct object is a full lexical NP that has undergone scrambling in the latter case, and a weak or clitic pronoun that has undergone some other kind of movement in the former cases.[16]

---

b. ??dass keiner [$_\alpha$ t_1 zu lesen]$_3$ versucht hat [$_{NP}$ das letzte Buch [$_{CP}$ das der Fritz darüber
that no-one to read tried has the last book that ART Fritz about that
verfaßt hat]]$_1$
written has

(i-a) is far from perfect, but as (i-b) shows, this fact is not related to remnant movement: If the infinitival occurs in situ, NP extraposition in this context has the same status.

Next, recall the hypothesis tentatively envisaged in section 7 of chapter 4, according to which string-vacuous scrambling in remnant movement constructions can be reanalyzed as an instance of extraposition. If nothing else is said, we are thus led to expect that there is a difference between sentences like (ii-a), which must involve two instances of scrambling, on the one hand, and sentences like (ii-b), which can involve a combination of scrambling and extraposition, and (ii-c), which must involve such a combination, on the other hand.

(ii) a. *Erneut hat [$_{VP}$ t_1 gerechnet]$_3$ [$_{PP}$ damit]$_1$ keiner t_3
again has counted therewith no-one

b. ??Erneut hat [$_{VP}$ t_1 gerechnet]$_3$ keiner t_3 [$_{PP}$ damit]$_1$
again has counted no-one therewith

c. ??dass [$_{VP}$ t_1 gerechnet]$_3$ keiner t_3 hat [$_{PP}$ damit]$_1$
that gerechnet no-one hat therewith

There are various potential intervening factors involved here, but by and large, this prediction seems to be borne out – (ii-b) and (ii-c) seem no worse than expected, given that they involve predicate scrambling and (the slightly marked phenomenon of) PP extraposition, and these examples are clearly less ill formed than (i-a). Thus, the data in (i-b) and (ii-b) & (ii-c) are compatible with the assumption that remnant scrambling is strictly ruled out only in case the antecedent of the unbound trace has also undergone scrambling, and lend further support to it. That said, it is not quite clear that these examples, given that they involve extraposition, actually have to exhibit an unbound trace t_1; cf. subsection 3.3.4 below.

[16]Note in passing that the data in (i-ab) exhibit the same contrast:

This observation can be replicated with coherent infinitives; see the contrast between (32-a) (= (21-c)) on the one hand (with the antecedent of the unbound trace located in a scrambling position), and (32-b) and (32-c) on the other (where the antecedent of the unbound trace is in a Wackernagel (-like) position):

(32) a. *[$_\alpha$ t$_1$ Zu lesen]$_,$ das hat [das Buch]$_1$ keiner t versucht
to read that has the book no-one tried

b. [$_\alpha$ t$_1$ Zu lesen]$_,$ das hat es$_1$ keiner t versucht
to read that has it no-one tried

c. [$_\alpha$ t$_1$ Zu lesen]$_,$ das hat's keiner t versucht
to read that has-it no-one tried

Thus, however these contrasts in remnant left dislocation are ultimately accounted for (see below), the conclusion that can again be drawn for the time being is that the preliminary generalization in (23) is wrong: It cannot be an inherent property of left dislocation not to apply to remnant XPs.

So far, I have argued against (23) by showing that scrambling and left dislocation can in principle affect remnant categories. Let me now turn to the complementary prediction made by (23), viz., that topicalization and *wh*-movement may systematically affect remnant categories (as long as no independent constraints like those discussed in the previous chapter are violated). Again, this prediction will turn out to be wrong.

3.3. Wh-Movement of Remnants is Sometimes Impossible

3.3.1. Some English Data

In section 2, we saw that *wh*-movement may apply to remnant categories (NPs and APs) in German; cf., e.g., (15-a) and (16), which are repeated here:

(33) a. [$_{NP}$ Was für ein Buch t$_1$]$_2$ hast du [$_{PP}$ über die Liebe]$_1$ t$_2$ gelesen ?
what for a book have you about the love read

(i) a. *[$_{VP}$ t$_1$ Gelesen]$_,$ das hat den Roman$_1$ keiner t
read that has the novel no-one

b. [$_{VP}$ t$_1$ Gelesen]$_,$ das hat (?)ihn/'n$_1$ keiner t
read that has him no-one

This shows that it is not a special property of the pronoun *es* and its reduced form *'s* to license otherwise impossible cases of remnant movement, but a more general property of weak (or clitic) pronoun movement. That said, it should be noted that certain instances of pronoun fronting do not fully succeed in lifting the ban on remnant left dislocation (or remnant scrambling, for that matter). In particular, remnant left dislocation and remnant scrambling appear to yield somewhat degraded results if the fronted pronoun is animate. I will leave open the question of what the nature of this apparent animacy effect is; cf. Cardinalletti & Starke (1995) for some relevant discussion.

b. [$_{AP}$ Wie stolz t$_1$]$_2$ ist der Fritz [$_{PP}$ auf sein Kind]$_1$ gestern t$_2$
how proud is ART Fritz of his child yesterday
gewesen ?
been

In contrast to that, Barss (1986, 57) and Saito (1989, 187) observe that *wh*-phrases may not be remnants in English. To see this, consider first some examples involving movement of complete *wh*-phrases from a *wh*-island, as in (34):

(34) a. ??[$_{NP}$ Which book about Nixon]$_2$ don't you know [$_{CP}$ whether to read t$_2$] ?
b. ?*[$_{NP}$ Which picture of John]$_2$ do you wonder [$_{CP}$ whether she likes t$_2$] ?

Wh-movement of a complete NP argument across a non-finite *wh*-island (as in (34-a)) gives rise only to a mild Subjacency-like effect in English (reducible to the Barriers Condition); see Chomsky (1986, 37), among others. *Wh*-movement of an argument from a finite *wh*-island (as in (34-b)) involves an additional tense effect that creates a somewhat stronger violation (see Frampton (1990, 69ff)). However, as noted by Barss and Saito, analogous examples involving *wh*-movement of a remnant *wh*-phrase, as in (35-a) (= (i) from footnote 20 of chapter 3) and (35-b), are severely ungrammatical in English, much more so than one would expect if only a Subjacency-like effect induced by an intervening *wh*-phrase were involved:

(35) a. *[$_{NP}$ Which book about t$_1$]$_2$ don't you know [$_{CP}$ who$_1$ to read t$_2$] ?
b. *[$_{NP}$ Which picture of t$_1$]$_2$ do you wonder [$_{CP}$ who$_1$ she likes t$_2$] ?

This effect is reminiscent of the situation found in German with scrambling and left dislocation that was discussed in section 2; accordingly, one might want to take the data in (35) to indicate that remnant *wh*-movement is impossible in English.[17] If this is so, we now of course have conflicting pieces of evidence – the English data in (35)

---

[17] As noted by Sigrid Beck (p.c.), one might try to rule out the sentences in (35) as involving semantically uninterpretable LFs. The basic idea would be that *who* in both cases is part of the restriction of the *wh*-operator *which* (interpreted as an existential quantifier) and consequently must be in the position of t$_1$ at LF; hence, it cannot show up in the SpecC position of the embedded clause. I will not take a firm stand on the issue of whether there can be an interpretable LF representation of (35-a) and (35-b) where *who* is in the embedded SpecC position, and what it could look like; this question is related to the well-known problem posed by sentences like (i-ab), where a *wh*-phrase is included in the restriction of another *wh*-phrase in a multiple question.

(i) a. [$_{NP}$ Which picture of [$_{NP}$ whom]] does she like ?
b. [$_{NP_1}$ Welches Buch von [$_{NP_2}$ welchem Autor]] hast du gelesen ?
which book by which author have you read

Rather, I would like to point out that an LF representation of, e.g., (35-a) that is a priori possible might look as in (ii), where (the copy) t$_1$ in (35-a) is interpreted as *who*, *who* in the embedded SpecC position is not interpreted at all, and the embedded *wh*-clause is understood as an *if*-clause:

(ii) [$_{NP}$ Which book about who]$_2$ don't you know [$_{CP}$ whether to read t$_2$] ?

appear to show that remnant *wh*-movement is impossible, whereas German data as in (33) seem to necessitate the assumption that remnant *wh*-movement is possible. Given such a situation, the first thing that comes to mind is that a parameter might be involved.

3.3.2. A Parameter?

Suppose that the contrast between the English and German data with respect to remnant movement is due to a parametric difference between these languages such that remnant *wh*-movement is simply never possible in English, whereas it is possible in German. If that is the case, we would expect German not to exhibit a contrast of the kind that shows up in (34) vs. (35) in English. However, although the German data are not as crystal-clear as one would want them to be (due to an intervening factor), this prediction turns out to be wrong.

Note first that the most striking contrast in the English data is that between (34-a) and (35-a), i.e., between the two cases in which the embedded *wh*-clause is non-finite. Unfortunately, however, there are no infinitival *wh*-clauses in German in the first place (see Tappe (1984), Giusti (1986), Wilder (1989), Lutz & Trissler (1992), and Sabel (1993), among others), so we cannot test whether a difference exists between remnant and non-remnant *wh*-extraction from infinitival *wh*-clauses or not. Thus, (36-a) (*wh*-movement of a complete NP across a non-finite *wh*-island) and (36-b) (remnant *wh*-movement across a non-finite *wh*-island) are indeed both ungrammatical, but this is independent of the issue of remnantness or even of long-distance *wh*-extraction, as shown by the illformedness of (36-c):

(36) a. *[$_{NP}$ Welches Buch über Benjamin]$_1$ weißt du nicht [$_{CP}$ ob PRO t_1
which book about Benjamin know you not whether
zu lesen] ?
to read

Thus, unless one can prove that (ii) cannot be derived from (35-a) by LF operations, a semantic explanation of the illformedness of the sentences in (35) does not seem possible. Furthermore, note that the same asymmetry shows up in cases where the *wh*-phrase in the embedded clause does not belong to the restriction of the *wh*-phrase of the matrix clause. Thus, *whose book* differs from *which book* in that *book* does not belong to the restriction of the *wh*-operator (rather, there is an implicit restriction "is a person"); cf., e.g., Heim (1992), Stechow (1993; 1996). With this in mind, consider (iii-ab):

(iii) a. ??[$_{NP}$ Whose book about John]$_2$ don't you know [$_{CP}$ whether to read t_2] ?
b. *[$_{NP}$ Whose book about t_1]$_2$ don't you know [$_{CP}$ who$_1$ to read t_2] ?

There is no semantic reason for *who* in (iii-b) to show up within NP_2 at LF (on the contrary, everything except *whose* must be reconstructed in NP_2); nevertheless, remnant *wh*-movement is impossible. To my mind, all this strongly suggests that a syntactic explanation for the illformedness of the sentences in (35) must be sought.

b. *[$_{NP}$ Welches Buch t$_2$]$_1$ weißt du nicht [$_{CP}$ [$_{PP}$ über wen]$_2$ PRO t$_1$ zu
which book know you not about whom to
lesen] ?
read

c. *Ich weiß nicht [$_{CP}$ [$_{PP}$ über wen]$_2$ PRO [$_{NP}$ ein Buch t$_2$]$_1$ zu lesen]
I know not about whom a book to read

Hence, the only way to test the hypothesis that a parameter is involved is to look at finite *wh*-clauses. Consider first (37):

(37) a. ?*[$_{NP}$ Welches Buch über die Liebe]$_2$ fragst du dich [$_{CP}$ ob du t$_2$
which book about the love ask you REFL whether you
lesen sollst] ?
read should

b. ?*[$_{NP}$ Was für ein Buch über die Liebe]$_2$ weißt du nicht [$_{CP}$ wann$_1$ du
what for a book about the love know you not when you
t$_1$ t$_2$ lesen sollst] ?
read should

These examples involve *wh*-extraction of a complete NP from a finite *wh*-island. There is some variation among speakers, but in general, the resulting sentences are considered ungrammatical to a fairly strong degree in German, perhaps even more so than their English counterpart in (34-b) (see subsection 4.2.5 of chapter 2 for relevant discussion of this effect). Still, though, a sentence in which a remnant *wh*-phrase is extracted across a *wh*-phrase in the embedded SpecC position that is the antecedent of the unbound trace in the *wh*-remnant is judged to be even more deviant by most speakers (the contrast is, of course, particularly striking with those speakers that find the examples in (37) tolerable); this indicates that an additional principle is violated.

(38) *[$_{NP}$ Was für ein Buch t$_1$]$_2$ fragst du dich [$_{CP}$ [$_{PP}$ über wen]$_1$ du t$_2$ lesen
what for a book ask you REFL about whom you read
sollst] ?
should

The contrast between (37-a)/(37-b) and (38) suggests that there is no parameter that distinguishes remnant movement in English from remnant movement in German – remnant *wh*-movement in German is also impossible if the antecedent of the unbound trace has undergone *wh*-movement.

This view is reinforced if we look at other Germanic languages, such as Danish. Violations of the *Wh*-Island Condition with *wh*-movement in the Scandinavian languages are somewhat weaker than in, say, German (see, e.g., Engdahl (1986)). Thus, *wh*-movement of an argument across a finite *wh*-island in a language like Danish is not impossible; this is shown in (39-a) (cf. Vikner (1995, ch. 2)). However, remnant *wh*-movement of the type in (38) in the same context is completely ungrammatical (cf. (39-b)), which strongly suggests an additional factor that is independent of *wh*-

islands.[18]

(39) a. ?[$_{\text{NP}}$ Hvad for en bog om kaerlighed]$_2$ ved du ikke [$_{\text{CP}}$ om du
what for a book about love know you not whether you
skal laese t_2] ?
should read

b. *[$_{\text{NP}}$ Hvad for en bog om t_1]$_2$ ved du ikke [$_{\text{CP}}$ [$_{\text{NP}}$ hvad]$_1$ du
what for a book about know you not what you
skal laese t_2] ?
should read

Returning now to German, we face a dilemma even within this language – data such as those in (33) seem to suggest that *wh*-movement may affect remnant categories in German, whereas data such as those in (38) seem to suggest that *wh*-movement may in fact not apply to remnant categories in the same language.

3.3.3. Long-Distance vs. Short Remnant Movement?

Since it does not seem practicable to assume that we are dealing with a case of parametric variation in the realm of remnant *wh*-movement here, let me pursue another hypothesis. This hypothesis rests on the idea that the well-formed remnant *wh*-movement data in (33) differ from the ill-formed remnant *wh*-movement data in (35), (38), and (39-b) in one respect – the *wh*-remnant has undergone short, i.e., clause-bound movement in (33), and long-distance movement in (35), (38), and (39-b). In fact, it turns out that almost all of the legitimate examples of remnant movement discussed up to now share the property that the remnant XP has not undergone long-distance movement from a finite clause. Thus, one might want to postulate a constraint on remnant movement, to the effect that remnant categories may never be separated from the antecedent of the unbound trace by an intervening finite clause boundary. The unbound trace, under this view, may not be "too far away" from its (non-c-commanding) antecedent. However, such a solution does not work either. As shown in (40), successive-cyclic long-distance *wh*-movement of remnant categories is possible in German:

(40) [$_{\text{NP}}$ Was für ein Buch t_1]$_2$ denkst du [$_{\text{CP}}$ t_2' dass Antje [$_{\text{PP}}$ über die Liebe]$_1$
what for a book think you that Antje about the love
t_2 gelesen hat] ?
read has

Similarly, successive-cyclic long-distance topicalization of remnants is permitted in German. Consider long-distance topicalization of a coherent infinitive in (41-a) (cf. Grewendorf & Sabel (1994, 303)), of a remnant NP in (41-b), and of a remnant VP in

[18] I am grateful to Sten Vikner (p.c.) for making me aware of this and providing the examples.

(41-c) (cf. subsection 7.3 of chapter 4):

(41) a. [$_\alpha$ t_1 Zu lesen]$_3$ glaubte sie [$_{CP}$ t'_3 habe keiner [$_{NP}$ das Buch]$_1$ t_3
to read believed she has$_{subj}$ no-one the book
versucht]
tried

b. [$_{NP}$ Ein Buch t_1]$_2$ meinte ich [$_{CP}$ t'_2 dass der Fritz [$_{PP}$ über die
a book thought I that ART Fritz about the
Liebe]$_1$ t_2 gelesen hat]
love read has

c. [$_{VP}$ t_2 Gerechnet]$_3$ denke ich [$_{CP}$ t'_3 dass [$_{PP}$ damit]$_2$ keiner t_3 hat
counted think I that therewith no-one has

Thus, the generalization that remnant movement is always clause-bound is wrong and cannot be used to account for the contrast between (33) on the one hand and (35), (38), and (39-b) on the other.[19]

To sum up so far, it is highly unlikely that variation in the domain of remnant *wh*-movement can be accounted for by either a parameter or a statement that requires remnant movement to be clause-bound. However, recall now the generalization that was arrived at in the case of remnant scrambling – this operation is illegitimate only if the antecedent of the unbound trace has also undergone scrambling. The very same generalization now appears to be descriptively correct with *wh*-movement: In the illicit cases of remnant *wh*-movement (cf. (35), (38), and (39-b)), the antecedent of the unbound trace has also undergone *wh*-movement; in the well-formed cases (cf. (33), and also (i) from the last footnote), the antecedent has undergone some other movement operation, viz., scrambling. Since scrambling does not exist in English, it therefore does not come as a surprise that data such as (33) do not show up and that the assumption that *wh*-movement may not apply to remnant categories in this language is generally taken to be uncontroversial. However, whereas English does not have scrambling, i.e., syntactic left-adjunction to XP, it has extraposition, which I assume here to be a case of syntactic right-adjunction to XP (see the previous chapter).

---

[19]Indeed, it is not even the case that remnant *wh*-movement is per se more sensitive to *wh*-islands than ordinary *wh*-movement. Consider a sentence like (i):

(i) ?*[$_{NP}$ Was für ein Buch t_1]$_2$ fragst du dich [$_{CP}$ ob du [$_{PP}$ über die Liebe]$_1$ t_2 lesen
what for a book ask you REFL whether you about the love read
sollst] ?
should

It appears that for those speakers that permit *wh*-extraction from a finite *wh*-island in German at all, (i) is more on a par with (37-a) than with (38).

3.3.4. Remnant Wh-Movement and Extraposition

If the hypothesis just formulated is correct and remnant *wh*-movement is systematically ruled out only if the antecedent of the unbound trace has also undergone *wh*-movement, we expect those instances of remnant *wh*-movement in English in which the antecedent of the unbound trace has undergone extraposition to be fine. This issue is discussed in Williams (1974, ch. 4) and Guéron & May (1984); a relevant example is given in (42):

(42) $[_{\text{NP}}$ How many pictures t_1 $]_2$ did you see t_2 $[_{\text{CP}_1}$ that I took yesterday] ?

Suppose that the extraposed CP is in this position as a result of syntactic movement (extraction from NP and right-adjunction to XP). The question then is: In what position is the extraposed CP exactly, i.e., can it bind its trace t_1 that is included in NP_2 in SpecC? In chapter 4, I argued that extraposition may end up in any XP-adjoined position. For (42) this means in a VP-adjoined, an IP-adjoined, or a CP-adjoined position (abstracting away from a more fine-grained structure of IP). If the order of rule application is *CP_1 extraposition* $\succ$ *wh-movement of NP_2*, CP_1 can be in any of these positions.[20] If, on the other hand, the order of rule application is *wh-movement of NP_2* $\succ$ *CP_1 extraposition*, then CP_1 must show up in the CP-right-adjoined position because of the Strict Cycle Condition.[21] This means that there is a structural ambiguity as to the position of CP_1 in (42); whether an *unbound* CP_1 trace shows up within the fronted NP_2 or not is unclear. This ambiguity does not yet disappear in a sentence like (43):

(43) $[_{\text{NP}}$ How many pictures t_1 $]_2$ do you think $[_{\text{CP}}$ t'_2 that John gave t_2 to Mary] $[_{\text{CP}_1}$ that I took yesterday] ?

Here, the *wh*-phrase NP_2 from which extraposition has applied shows up in the SpecC position of the matrix clause; but again, there are two options. One possibility is that CP_1 is extraposed before *wh*-movement of NP_2. Since extraposition must always be clause-bound (the Upward Boundedness Constraint; cf. (110) from chapter 4), in this case subsequent *wh*-movement must be remnant movement, creating an unbound

[20]Note that even if CP_1 is right-adjoined to CP, subsequent *wh*-movement will not violate the Strict Cycle Condition, as assumed here (cf. (52) from chapter 3). The reason is that both operations end up in the same minimal residue, viz., that of the matrix C node.

[21]Given what has been said in chapter 4, one might assume that this latter derivation is not possible because extraction from NP_2 would then take place at a point in the derivation at which NP_2 is no longer in a selected position and therefore qualifies as a barrier. However, as was noted in footnote 13 of chapter 4, it seems that extraposition may freely escape from NPs even if these NPs are barriers in their in situ positions (e.g., in the case of subjects), in contrast to leftward movement types. Thus, regardless what explanation of this fact ultimately proves to be correct (see Baltin (1983), Culicover & Rochemont (1990), and Müller (1996)), there is a priori no reason why extraposition should not apply from an NP in SpecC, given that extraposition has a way of circumventing NP barriers that is not available for other movement types.

trace. The other possibility, though, is that first NP_2 undergoes long-distance *wh*-movement, and then CP_1 extraposition takes place in the matrix clause, right-adjoining CP_1 to the matrix CP, an option that seems real (see the last footnote); in that case, (43) could be analyzed without remnant *wh*-movement. It seems that examples like (44) avoid the structural ambiguity:

(44) $[_{NP}$ Which pictures t_1 $]_2$ do you think $[_{CP}$ t'_2 that no-one gave t_2 to Mary $[_{PP_1}$ of any of her friends]] ?

Suppose that licensing of polarity items like *any* depends on a combination of precedence and m-command in overt syntax (see Barss & Lasnik (1986) and Jackendoff (1990)). Then, *any* in (44) must be in the same clause as *no-one* at S-structure (Spell-Out); i.e., it must be in the most deeply embedded CP. But if this is so, *wh*-movement of NP_2 to the matrix SpecC position must take place later and must therefore involve remnant movement.

Much more would have to be said about this issue, and one might possibly make the point that the argument for remnant *wh*-movement in extraposition structures in English is not decisive.[22] However, it is at least worth bearing in mind that there is an obvious difference in English between illicit traces in *wh*-phrases in examples like (35) (where the antecedent has undergone *wh*-movement, too) and licit traces in *wh*-phrases in examples like (42), (43), and (44) (where the antecedent has undergone extraposition); and this difference corresponds exactly to our expectations, given that remnant *wh*-movement is strictly blocked only if the antecedent of the unbound trace has also been *wh*-moved (and not if it has undergone some other kind of movement).[23]

That said, let me now turn to the last kind of remnant movement discussed in section 2, viz., topicalization, and see whether the same effect shows up as with *wh*-movement, scrambling, and left dislocation.

---

[22] Note in passing that the English facts discussed in this subsection find close parallels in German.

[23] In fact, the analogy between remnant *wh*-movement & scrambling in German and remnant *wh*-movement & extraposition in English goes even further. It seems that remnant *wh*-movement in English may indeed apply across a (non-finite or finite) *wh*-island as long as the antecedent of the NP-internal trace is not *wh*-moved but extraposed. This is shown by the data in (i-a) (non-finite *wh*-island) and (i-b) (finite *wh*-island) (Steve Abney & Chris Wilder (p.c.)):

(i) a. ??$[_{NP}$ How many pictures t_1 $]_2$ do you wonder $[_{CP}$ whether to give t_2 to Mary $[_{CP}$ that I took yesterday $]_1$] ?

b. ?*I would love to know $[_{CP}$ $[_{NP}$ how many pictures t_1 $]_2$ he wondered $[_{CP}$ whether he should show t_2 to his mother $[_{CP}$ that he took yesterday $]_1$]]

(i-ab) are no worse than their counterparts in which CP extraposition has not applied (i.e., there is no violation of a specific constraint on remnant movement in addition to the standard *wh*-island effect, which can be reduced to the Barriers Condition), and they are completely analogous to the German example (i) in footnote 19, where a combination of scrambling from NP and *wh*-movement of NP across a *wh*-island occurs.

3.4. Remnant Topicalization is Sometimes Impossible

In all the well-formed examples involving remnant topicalization that have been discussed so far, the antecedent of the unbound trace has undergone scrambling. Given the conclusions of the preceding subsections, it would not be surprising if remnant topicalization were impossible if the antecedent of the unbound trace had also undergone topicalization. And indeed, an example like (45) is completely ungrammatical:

(45) *[$_{\alpha}$ t_1 Zu lesen]$_2$ glaube ich [$_{\text{TopP}}$ [$_{\text{NP}}$ dieses Buch]$_1$ hat keiner t_2 versucht]
to read think I this book has no-one tried

Given the present system of assumptions (more precisely, the Barriers Condition, the Strict Cycle Condition, and the economy constraints Fewest Steps and Last Resort; see chapter 4), (45) must be derived by NP_1 topicalization (out of the coherent infinitive) followed by α-topicalization across the embedded topic NP_1 to the matrix SpecTop position. The illformedness of the resulting sentence at first glance seems to show that the same effect is at work as with the other movement types that were discussed in the preceding subsections. Unfortunately, though, there is an intervening factor with topicalization as in (45) that makes it impossible to determine whether this effect does indeed occur or not, i.e., whether there is a special constraint that bans remnant topicalization if the antecedent of the unbound trace has also undergone topicalization.

The intervening factor is the following: As was already noted in footnote 48 of chapter 4, it seems to be a robust fact about the Germanic languages that embedded topicalization induces a strict island for long-distance movement of any kind (cf. Zaenen (1980), den Besten (1989), Rochemont (1989), Vikner (1990; 1995), and Müller & Sternefeld (1993), among many others). Thus, consider the contrast in German between successive-cyclic topicalization in (46-a) and topicalization across an embedded topic in (46-b):

(46) a. [$_{\text{NP}}$ Dieses Buch]$_1$ glaube ich [$_{\text{TopP}}$ t_1' sollte man t_1 lesen]
this book think I should one read

b. *[$_{\text{NP}}$ Dieses Buch]$_1$ glaube ich [$_{\text{TopP}}$ man$_2$ sollte t_2 t_1 lesen]
this book think I one should read

Topicalization of a direct object as in (46-b) is far more ungrammatical than one would expect it to be if only a Subjacency-like violation were involved. Thus, regardless of what eventually accounts for the strict nature of topic islands (see the references just cited for proposals), it is evident that remnant topicalization as in (45) is ruled out in the same way. Accordingly, if there is a special prohibition against remnant topicalization in this context, it is blurred by the more general constraint that rules out any movement across an embedded topic, and therefore cannot be directly attested, unlike what was the case with *wh*-movement. Ultimately, this is so because topic islands, in contrast to *wh*-islands, are always strict and block extraction of any kind, irrespective of the properties of the moved item. On the other hand, this means that in the absence of any evidence to the contrary, we can safely assume that there is indeed a close parallelism between topicalization and other movement types (scrambling, *wh*-movement) and that an example like (45) would indeed be ungrammatical even if embedded topics did not turn the minimal clause they show up in into a strict island. Based on this, let me now turn to what I take to be the correct generalization underlying the data in sections 2 and 3.

3.5. Conclusion

The discussion of the preceding subsections has revealed that it is not an inherent property of a movement type to apply or not to apply to remnant XPs and create unbound traces; so, the preliminary generalization (23) is proved false. Rather, it seems that remnant scrambling is strictly prohibited only if the antecedent of the unbound trace has also undergone scrambling, and otherwise possible (if no other constraints of grammar are violated). Similarly, remnant *wh*-movement turned out to be illegitimate if the antecedent of the unbound trace has also undergone *wh*-movement, but not if the antecedent has undergone some other kind of movement. Thus, what we can minimally conclude from this is that whether or not remnant movement is possible does not depend only on (i) the position of the remnant XP, but also (ii) on the position of the antecedent of the unbound trace. More specifically, I would like to contend that the correct descriptive generalization covering all the relevant data from this and the previous section is (47), where "Y-movement" stands for *wh*-movement, scrambling, topicalization, etc.[24]

(47) *Generalization* (final version):
Remnant XPs cannot undergo Y-movement if the antecedent of the unbound trace has also undergone Y-movement.

According to (47), there is a kind of non-identity requirement on remnant movement: Remnant XPs cannot undergo a certain type of movement if the antecedent of the unbound trace has undergone the same type of movement. Clearly, (47) implies that the topicalization example (45) is excluded irrespective of the strong islandhood of clauses with internal topicalization; it also presupposes that left dislocation and scrambling count as the same movement type for the purposes of (47).

Assuming that (47) is empirically correct, I will now try to derive this generalization.

[24] Also see Takano (1993, 14), who has independently suggested a similar generalization.

4. Unambiguous Domination

4.1. The Principle

The first thing to note is that the generalization in (47) does not yet follow from any of the constraints on movement (locality, economy, and cyclicity) assumed so far; so, it appears as though a new constraint is called for. Given the generalization in (47), I would like to propose the principle in (48), which will serve as a constraint on remnant movement.

(48) *Unambiguous Domination*:
An α-trace must not be α-dominated.

Suppose that "α-trace" means "trace with a (not necessarily c-commanding) antecedent in a position of type α" and that "α-dominated" means "dominated by a category in a position of type α." (48) then predicts that a remnant XP may not occupy a position of a certain type if it contains a trace that has an antecedent in the same type of position. As for the set of positions that α ranges over, I will assume, in line with the discussion in chapter 2 and afterwards, that (48) at least distinguishes between the landing sites in (49):

(49) a. SpecC: The landing site of *wh*-movement at S-structure.
b. SpecTop: The landing site of topicalization.
c. XP-Left-Adj: The landing site of scrambling (X = V or I in German) and left dislocation (X = C) (see below).
d. XP-Right-Adj: The landing site of extraposition.
e. SpecF and F-Adj: The landing sites of weak pronoun movement and Wackernagel-cliticization in German (the nature of F is left open).
f. SpecX: The landing site of (Case-driven) A-movement (X = I or V (see below), perhaps AGR_O in some languages)

(48) then amounts to a requirement that if extraction from an XP has taken place, ending up in any of the six positions mentioned in (49),[25] the remnant XP itself can only be moved to a position of one of the other five types.[26]

---

[25]Or rather, seven positions, given that weak pronoun movement and pronominal cliticization must be formally distinguished as indicated – a matter which, however, I do not take to be settled.

[26]At this point, recall again the assumption made in chapter 2 (footnote 15), to the effect that AGR_S, T, and AGR_O are merged into a single I node in German. If there were Case-driven object movement to SpecAGR_O in German, and if this movement could apply at S-structure (as is sometimes argued), then there would be a problem for Unambiguous Domination – we would have to ensure that overt movement to SpecAGR_O could not create representations that I have so far attributed to the effect of scrambling; otherwise, an account of illicit cases of remnant scrambling in terms of Unambiguous Domination would potentially be undermined (since the two positions in question would not be identical). Hence, the assumption that Case-driven movement to SpecAGR_O does not exist in German is corroborated by the existence

Before I turn to an illustration of how (48) rules out illicit cases of remnant movement, the questions of what kind of constraint (48) is and where it applies must be answered. As it stands, the Unambiguous Domination requirement is a representational constraint, so the only domains of application that are a priori conceivable are S-structure and LF. At first glance, one might think that an LF application of (48) is out of the question, the reason being that LF movements like quantifier raising and *wh*-movement may destroy S-structure representations that exhibit remnant movement configurations that violate the Unambiguous Domination requirement. However, recall from sections 4 and 6 of chapter 3 that I have adopted the copy theory of movement, and that in a particularly strong form: Every instance of movement leaves a complete copy as a trace, and although copies (and parts of copies) are often not interpreted at LF, neither a full copy nor a part of one is actually deleted at LF. A consequence of this assumption (that was made in chapter 3 in order to deal with the fact that overt movement that is reconstructed is not blocked by Fewest Steps) is that no S-structure information gets lost on the way to LF. So, if there is a trace with an antecedent in, say, a scrambling position (e.g., in VP-Adj) that is dominated by a category (e.g., an NP) in another scrambling position (e.g., in IP-Adj), thus violating (48) at S-structure, then no LF operation will be able to remove that illicit configuration, and (48) will still be violated at LF. For this reason, I will assume that we can dispense with the assumption that Unambiguous Domination is an S-structure constraint and can take it to apply solely at LF. This is, of course, in line with the leading idea of the minimalist program that there are no constraints referring to S-structure and that S-structure itself is superfluous as a syntactic level of representation.[27]

Given these assumptions about landing sites and the domain of application, it remains to be shown that the Unambiguous Domination requirement (48) makes the correct predictions for the data discussed so far.

---

of illicit cases of remnant scrambling, under present assumptions.

[27] A few remarks are in order. First, note that (48) can be straightforwardly reformulated as a derivational constraint, i.e., as a constraint on movement, not on (LF) representations; cf. (i).

(i) *Unambiguous Domination* (derivational version):
In a structure ... $[_A$ ... B ...] ..., A and B may not undergo the same kind of movement.

The effect of (i) is basically identical to that of (48). I will nevertheless maintain the representational version of Unambiguous Domination in what follows; one should keep in mind, however, that nothing really hinges on this for the issues currently under discussion; but see section 7 below. Second, (48) is clearly too strong as it stands, since it generally precludes α-domination of an α-trace, no matter how deeply embedded the α-trace is in another category located in an α-position. (Similar conclusions apply in the case of (i).) In section 6, I will show that the requirement that an α-trace must not be α-dominated holds only in a certain local domain. Finally, (48) applies to all kinds of traces – bound and unbound, variables, anaphors, and X^0 traces. Presently, however, I am only concerned with unbound variables; the relevance of (48) for other kinds of traces is discussed in section 5 below.

4.2. Remnant Scrambling

Consider first scrambling of remnant categories, as depicted schematically in (50):

(50) ... $[_{XP}$ $[_{\gamma}$ ... t_1 ...] $[_{XP}$ ... YP_1 ...]] ...

By assumption, the remnant category γ is adjoined to VP or IP. Suppose first that YP has also undergone scrambling and thus occupies an adjunction position. (48) then requires that t_1, the trace of YP, must not be dominated by a category γ in a scrambling position. This rules out remnant scrambling in all the illicit cases we have seen so far (viz., scrambling of coherent infinitives in (2-a) and (2-b), scrambling of remnant VPs in (6-a) – (6-c), scrambling of remnant APs in (9-a) and (9-b), and scrambling of remnant NPs in (12) and (13)).

In contrast, suppose now that YP in (50) has not also undergone scrambling, but some other kind of movement, like, e.g., weak pronoun movement or cliticization, as in the examples in (29), (30-b), and (30-c). In this case, (48) does not rule out remnant scrambling – a trace of weak pronoun movement (or cliticization) is dominated by a remnant XP in a scrambling position, and the sentences are correctly predicted to be well formed (since no other constraint is violated).

Next, note that we would in principle also expect cases of remnant scrambling to be possible where the unbound trace is a trace left by some other movement type, like topicalization or *wh*-movement. But since scrambling is clause-bound in languages like German, and since there is in general no clause-internal adjunction site that would precede the landing site of either topicalization or *wh*-movement, the relevant constructions are all independently excluded. For instance, (48) does not per se rule out examples like those in (51), because here t_1 is dominated by an XP in an adjoined position, whereas the antecedent of t_1 occupies a SpecC position:

(51) a. *dass $[_{\alpha}$ t_1 zu lesen $]_2$ ich nicht weiß $[_{CP}$ [welches Buch $]_1$ ich t_2
that to read I not know which book I
versuchen sollte]
try should

b. *Ich weiß nicht $[_{CP}$ $[_{\alpha}$ t_1 zu lesen $]_2$ $[_{CP}$ [welches Buch $]_1$ ich t_2
I know not to read which book I
versuchen sollte]]
try should

However, (51-a) involves long-distance scrambling from a finite clause and is thus ruled out by whatever accounts for the clause-boundedness of scrambling in German (see chapter 2).[28] On the other hand, remnant scrambling to the embedded CP, as in (51-b), is ruled out because CP does not qualify as an adjunction site in German (with

[28]Coherent infinitives, of course, permit long-distance scrambling; but infinitives in German do not have a landing site for either *wh*-movement or topicalization.

the exception of left dislocation structures; see below).[29]

4.3. Remnant Wh-Movement

Next consider *wh*-movement of remnants, as in (52):

(52) ... $[_{\text{CP}}$ $[_{\gamma}$ ... t_1 ...] $[_{\text{C'}}$ ... YP_1 ...]] ...

Here, the remnant category γ is in a SpecC position. If YP also occupies a SpecC position, (48) is violated and ungrammaticality results – t_1 is dominated by an XP in SpecC, and its antecedent also occupies a SpecC position. This correctly rules out (35-a) and (35-b) in English, (38) in German, and (39-b) in Danish. If, on the other hand, YP is not in a SpecC position in (52), but rather in a scrambling position, *wh*-movement of a remnant category is predicted to be possible – the Unambiguous Domination requirement (48) is not violated, because t_1 is dominated by an element in a SpecC position and its antecedent occupies a scrambling (i.e., left-adjunction) position. This accounts for the wellformedness of examples like (15-a) & (15-b) (*wh*-movement of remnant NPs), (16) (*wh*-movement of a remnant AP), (40) (successive-cyclic long-distance *wh*-movement of a remnant NP), and (i) of footnote 19 (where *wh*-movement of an NP applies from a *wh*-island) in German. Similarly, if YP in (52) is an extraposed item that does not bind its trace, as is arguably the case in the English examples in (42), (43), and (44), Unambiguous Domination is respected, and remnant *wh*-movement is possible.

4.4. Remnant Topicalization

Let me now turn to topicalization of remnant categories. Given that topicalization is movement to SpecTop, remnant topicalization typically involves configurations such as the one depicted in (53):

(53) ... $[_{\text{TopP}}$ $[_{\gamma}$ ... t_1 ...] $[_{\text{Top'}}$ ... YP_1 ...]] ...

Under the analysis of topicalization as substitution in SpecTop, we expect topicalization of remnants to be impossible if the antecedent of the unbound trace (YP in (53)) occupies a SpecTop position and possible otherwise (if no other conditions of grammar are violated). This prediction is borne out. Thus, we have seen that remnant topicalization as in (53) is allowed if YP has undergone scrambling; cf. topicalizati-

[29]Thus, examples like (i) that involve adjunction of a complete XP to CP in German are also strongly ungrammatical:

(i) *dass ich nicht weiß $[_{\text{CP}}$ $[_{\text{NP}}$ dieses Buch $]_1$ $[_{\text{CP}}$ warum$_2$ $[_{\text{IP}}$ ich t_2 t_1 lesen sollte]]]
that I not know this book$_{acc}$ why I read should

Therefore, the ungrammaticality of (51-b) follows irrespective of the issue of remnant movement.

on of coherent infinitives in (1-a) and (1-b), topicalization of remnant VPs in (5-a) – (5-c), topicalization of a remnant AP in (8), topicalization of remnant NPs in (11-a) and (11-b), and long-distance topicalization of coherent infinitives and remnant NPs in (41-a) and (41-b). In contrast, topicalization of remnants is predicted to be illicit if the antecedent of the unbound trace itself occupies a SpecTop position, as in (45). However, as noted before, this result is more or less vacuous, since the relevant construction is ruled out on independent grounds: As soon as YP in (53) is in SpecTop, it blocks movement of any item across this position, regardless of whether the item is a remnant or a complete XP.

Still, there is a bit more to be said about remnant topicalization. Given that the landing sites of *wh*-movement and topicalization differ, as I have assumed throughout, it follows that remnant topicalization should not violate (48) if the antecedent of the unbound trace is *wh*-moved rather than topicalized. Again, the prediction is confirmed. To see this, recall first from subsection 4.2.5 of chapter 2 that although *wh*-clauses in general block long-distance *wh*-movement in German, this is not the case with topicalization from a *wh*-clause – in this case, only a weak Subjacency-like violation occurs with object movement (cf. Fanselow (1987) and others):

(54) a. ??[$_\alpha$ Bücher zu lesen]$_1$ weiß ich nicht [$_{CP}$ warum sie t_1 versucht hat]
books to read know I not why she tried has

b. ??[$_{NP}$ Ein Buch über die Liebe]$_2$ weiß ich nicht [$_{CP}$ wieso Antje t_2
a book about the love know I not why Antje
gekauft hat]
bought has

Interestingly, remnant topicalization as in (55-a) & (55-b) is neither better nor worse than the examples in (54) (cf. also den Besten & Webelhuth (1990) and Fanselow (1992)):

(55) a. ??[α t_1 Zu lesen]$_2$ weiß ich nicht [$_{CP}$ was$_1$ sie t_2 versucht hat]
to read know I not what she tried has

b. ??[NP Ein Buch t_1]$_2$ weiß ich schon [$_{CP}$ worüber$_1$ Antje t_2 lesen sollte]
a book know I PRT about-what Antje read should

Given Unambiguous Domination in (48), this is to be expected. Topicalization of a coherent infinitive in (55-a) and of a remnant NP in (55-b) does not violate Unambiguous Domination because the unbound trace included in the topicalized items has an antecedent which occupies SpecC (and not SpecTop).[30]

---

[30] As observed by Longobardi (1985), examples like (i) are ungrammatical in Italian:

(i) *[$_\alpha$ A sposare t_1]$_2$ non so proprio [$_{CP}$ [$_{NP}$ quale ragazza]$_1$ Gianni sarebbe disposto t_2]
to marry I really do not know which girl Gianni would be ready

4.5. Remnant Left Dislocation

Finally, let me turn to left dislocation. Recall from the discussion of the examples in (19) and (21) that left dislocation of remnants is impossible if the antecedent of the unbound trace has undergone scrambling, i.e., adjunction. This result follows from (48) if we assume that left-dislocated XPs occupy a left-adjoined position, so that, in effect, left dislocation is formally indistinguishable from scrambling. I would like to suggest that left dislocation in German as in (56) involves the marked option of left-adjunction to CP, licit only because of co-indexing with a pronoun in the SpecC position (recall that CP is not a possible landing site for scrambling in German):

(56) (dass ich glaube) [$_{CP}$ den Fritz$_1$ [$_{CP}$ den$_1$ hat keiner t$_1$ gesehen]]
that I believe ART Fritz him has no-one seen

Under this assumption, a typical remnant left dislocation structure looks as in (57):[31]

(57) ... [$_{CP}$ [$_{\gamma_2}$ ... t$_1$ ...] [$_{CP}$ pronoun$_2$ ... YP$_1$ ... t$_2$ ...]] ...

Evidently, (i) closely resembles the German examples in (55) and hence should not violate Unambiguous Domination under current assumptions. However, the illformedness of (i) is accounted for if we assume that topicalization in Italian is to be analyzed like *wh*-movement (i.e., as movement to SpecC) and does not count as a movement type sui generis. (One might speculate that TopP is either completely absent here or merged with CP, the idea being that T is too "weak" to project a TopP on its own in the non-V/2 language Italian – note that in Müller & Sternefeld (1993) the existence of TopP in the Germanic languages is closely tied to the occurrence of V/2 phenomena.) Thus, if topicalization and *wh*-movement are formally indistinguishable in Italian, an α-trace is α-dominated in (i), in conflict with Unambiguous Domination. A similar conclusion might be drawn on the basis of data like (ii-a) in English:

(ii) a. *[$_{\alpha}$ Ready to marry t$_1$]$_2$ I wonder [$_{CP}$ who$_1$ John is t$_2$]
b. ??[$_{\alpha}$ Give t$_1$ to Mary]$_2$ I really don't know [$_{CP}$ [$_{NP}$ which book]$_1$ he did t$_2$]

However, examples like (ii-b) appear to behave more like topicalizations of complete VPs across a finite *wh*-island in English and thus suggest that only a Subjacency-like *wh*-island violation is involved with remnant VP topicalization in this language. In the present framework, this would imply that Unambiguous Domination can be met in (ii) after all and that (ii-a) must be explained differently. I will have to leave this problem open.

[31] Note in passing that the analysis in this subsection does not force a decision as to whether left dislocation is to be analyzed in terms of (exceptional) movement of the left-dislocated item or via base-generation – i.e., whether t$_2$ is a trace of γ_2 or (only) of the co-indexed pronoun. For the sake of concreteness, I will assume the former. It goes without saying that left dislocation raises a number for further questions that I will have to leave unanswered in this study. To name just a few: It is unclear why exactly a co-indexed pronoun lifts the ban on left-adjunction to CP in German, and why only a co-indexed pronoun can do so (and not some arbitrary other item). Furthermore, an articulated analysis of left dislocation would have to clarify the question of whether the pronoun in SpecC is the partial Spell-Out of a chain-internal copy trace, or a resumptive element that is inserted after the movement operation has taken place. See Chomsky (1977), Cinque (1977; 1990), Koster (1978, 199ff), and Cardinaletti (1986) for relevant discussion. Also, see section 7 below, where the movement analysis of left dislocation does in fact become relevant.

Now, if the left-dislocated XP γ in (57) occupies a left-adjoined position, it follows that left dislocation behaves exactly like scrambling with respect to remnants. Thus, the examples in (19) and (21) are now excluded – if the antecedent of the unbound trace YP occupies a scrambling position, the trace may not be dominated by a category in a left-adjoined position due to the Unambiguous Domination requirement.[32] On the other hand, if YP occupies another kind of position, as in the grammatical cases involving weak pronoun movement or pronominal cliticization that were discussed in subsection 3.2 (cf. (31-b), (31-c), (32-b), and (32-c)), Unambiguous Domination is respected, and grammaticality results, as desired.

Note next that left dislocation can apply long-distance in German:

(58) [$_\alpha$ Dieses Buch zu lesen]$_,$ das denke ich dass man t versuchen sollte
this book to read that think I that one try should

Given that long-distance left dislocation is possible, (48) makes an interesting prediction. Unbound traces in remnants which have undergone long-distance left dislocation should be able to take an antecedent in a SpecC position, but still not in a scrambling position. As expected, ungrammaticality results in the case of long-distance left dislocation if the antecedent of the unbound trace has undergone clause-bound scrambling; cf. (59):

(59) a. *[$_\alpha$ t_1 Zu lesen]$_,$ das denke ich dass man [$_{NP}$ dieses Buch]$_1$ t versuchen sollte
to read that think I that one this book try should

b. *[$_\alpha$ t_1 Zu lesen]$_,$ das weiß ich nicht warum man [$_{NP}$ dieses Buch]$_1$ t versuchen sollte
to read that know I not why one this book try should

In (59-a) the antecedent of the unbound trace occupies a scrambling position, and the trace t_1 included in the left-dislocated remnant XP violates principle (48): It is dominated by a category in a left-adjoined position and its antecedent is also located in a left-adjoined position. Basically, the same situation holds in (59-b) – here, left dislocation across a *wh*-island has applied, t_1 again violates (48), and the sentence

[32]To account for the Dutch data reported in footnote 7, we would have to motivate the hypothesis that the position occupied by γ_2 in (57) counts as different from scrambling positions – an enterprise that does not strike me as necessarily futile in light of the fact that Dutch scrambling behaves differently from German scrambling in a number of respects (e.g., as noted in footnote 12 of chapter 1, Dutch scrambling is strictly order-preserving, i.e., argument NPs are not actually subject to permutation). Thus, if Dutch scrambling were movement to designated specifier positions, as has sometimes been argued, while left dislocation would still involve left-adjunction to CP, we would expect that remnant left dislocation is much more liberal than remnant scrambling in this language.

is strongly ungrammatical, more so than one would expect if only a Subjacency-like effect were involved. (Recall from the discussion of (54) that extraction of a [–wh] element from a *wh*-island in German typically results in a mild violation.)

Now, given the option of left dislocation applying long-distance in German, it seems that remnant left dislocation should in principle be possible if the antecedent of the unbound trace is located in a non-scrambling position like SpecC. Here, the data are rather murky, and judgements are not very clear. But it seems that examples like (60) are indeed better than those in (59), which follows directly from the assumptions made so far:

(60) ?*[$_\alpha$ t$_1$ Zu lesen]$_,$ das weiß ich nicht was$_1$ man t versuchen sollte
to read that know I not what one try should

Still, examples like (60) strike speakers as somewhat worse than analogous sentences involving topicalization, as in (55-a) and (55-b). This, at first glance, is unexpected – Unambiguous Domination is found in (60) as it is in (55). However, it appears that an independent factor intervenes in (60), as compared to (55). (61) shows that left dislocation across a *wh*-island is in general worse than the corresponding topicalization case for some reason (which is not clear to me).

(61) ?*[$_\alpha$ Dieses Buch zu lesen]$_,$ das weiß ich nicht warum man t versuchen sollte
this book to read that know I not why one try should

Thus, all in all, the predictions made by the Unambiguous Domination requirement (48) for remnant movement are confirmed by long-distance left dislocation.

4.6. Conclusion

Summarizing so far, I have argued that the principle of Unambiguous Domination, as formulated in (48), explains remnant movement asymmetries in a relatively straightforward way. Thus, it has turned out that what at first looked like a peculiar movement restriction on scrambling (viz., the impossibility of moving a remnant XP) does not have to be accounted for by invoking construction-specific assumptions; on the contrary, the ban on remnant scrambling has been shown to have parallels in other movement types – there is evidence that *wh*-movement and left dislocation may not apply to remnants either under certain circumstances, which are predicted correctly by the condition (48).

It seems that one could stop here. Nevertheless, I think that more must be said about the empirical range and the conceptual status of the principle of Unambiguous Domination in (48). In particular, three remaining questions should be answered, viz.:

1. Does the Unambiguous Domination requirement apply to all kinds of traces (unbound and bound, variables, anaphors, and X^0 traces)?

2. In which local domain does the Unambiguous Domination requirement hold?

3. Why does the Unambiguous Domination requirement hold?

I will address these questions in turn in the following three sections.

5. Trace Types and Unambiguous Domination

First, let us try to figure out exactly which kinds of traces obey the Unambiguous Domination requirement (48). As it stands, this condition applies not only to unbound traces but also to bound traces. However, so far we have only encountered positive evidence for Unambiguous Domination with respect to unbound traces (in other words, with remnant movement). It remains to be shown that Unambiguous Domination does not produce unwanted consequences for the distribution of bound traces. Furthermore, all the traces I have been concerned with so far are traces of A-bar movement, i.e., variables. Again, however, (48) is formulated in a more general way and also covers traces of A-movement (anaphors) and traces of head movement (X^0 traces). Let me first turn to the issue of bound vs. unbound traces, next treat the problem of variables vs. anaphors, and then turn to X^0 traces.

5.1. Bound vs. Unbound Traces

Suppose that the Unambiguous Domination requirement (48) does indeed apply to bound traces, too; this clearly is the null hypothesis. The principle then implies that bound scrambling traces may not be contained in scrambled XPs, that bound traces of topicalization may not occur in topicalized XPs, that bound traces of *wh*-movement cannot be dominated by *wh*-phrases in SpecC, and so forth. These predictions are indeed borne out. Consider, for instance, German examples like (62-a) and (62-b):

(62) a. *dass [$_{NP}$ den Aufsatz]$_1$ mal wieder [$_{VP}$ t$_1$ gelesen]$_2$ keiner t$_2$ hat
that the article$_{acc}$ once again read no-one has

b. *dass darüber$_1$ gestern [$_{NP}$ ein Buch t$_1$]$_2$ keiner t$_2$ gelesen hat
that about that yesterday a book$_{acc}$ no-one read has

In (62-a) an NP has been scrambled out of a VP which has itself been scrambled, and the result is completely ungrammatical; in particular, it is far more deviant than we would expect if only the general prohibition against predicate scrambling mentioned above were involved. In (62-b) a PP is scrambled from an NP which has itself undergone scrambling to a position in front of the subject. Since PP scrambling from NP is possible with a verb like *lesen* (‘read’) if the NP is located in situ, the ungrammaticality of (62-b) must be due to the fact that the NP has undergone scrambling.

In the same way, topicalization from a topicalized NP or VP is illicit in German, as shown by the examples in (63):

(63) a. *[$_{PP}$ Über die Liebe]$_1$ glaube ich [$_{NP}$ ein Buch t$_1$]$_2$ hat Antje t$_2$ gelesen
about the love believe I a book has Antje read

b. *[$_{NP}$ Ein Buch]$_1$ denke ich [$_{VP}$ dem Peter t_1 gegeben]$_2$ hat die
   a book$_{acc}$ think I ART Peter$_{dat}$ given has ART
   Claudia t_2
   Claudia$_{nom}$

Next, *wh*-movement from an NP which has itself been *wh*-moved is impossible in German ((64-a) = (13-b) from chapter 4):[33]

(64) a. *Worüber$_1$ hast du gesagt [$_{CP}$ [$_{NP}$ was für Bücher t_1]$_2$ er t_2 gelesen
   about what have you said what for books he read
   hat] ?
   has

b. *[$_{PP}$ Über wen]$_1$ fragst du dich [$_{NP}$ was für ein Buch t_1]$_2$ du t_2
   about whom ask you REFL what for a book you
   lesen solltest ?
   read should

And finally, left-adjunction (i.e., scrambling) from a VP which has undergone left dislocation results in strong ungrammaticality in German. This is shown in (65):

(65) *dass er [$_{NP}$ das Buch]$_1$ sagt [$_{VP}$ t_1 gelesen], das habe t keiner
   that he the book$_{acc}$ says read that has$_{subj}$ no-one

In the examples in (62) through (65), Unambiguous Domination is violated, and the sentences are indeed ungrammatical. Thus, it appears that the assumption that bound traces must be dominated unambiguously does not produce unwanted consequences.[34] On the other hand, it is a fact that most of the configurations now excluded by Unambiguous Domination are ill formed anyway because they involve Freezing effects. Thus, they receive an account that is identical to that given for examples such as those in (66) (cf. the data in subsection 2.1 of chapter 4):

(66) a. *Worüber$_1$ hat [$_{IP}$ [$_{NP}$ ein Buch t_1]$_2$ [$_{IP}$ keiner t_2 gelesen]] ?
   about-what has a book$_{acc}$ no-one read

---

[33] That there is no difference between bound and unbound *wh*-traces in this respect in German is also noted by Grewendorf (1994, 41).

[34] It seems that a bit more would have to be said about Japanese, though. In this language, unbound scrambling traces in scrambled items are not permitted, just as in German (cf. subsection 2.1 above). However, there are contexts in which bound scrambling traces in scrambled items are permitted, in violation of Unambiguous Domination. In what follows, I will abstract away from this apparent instance of language-specific variation. See, however, Sauerland (1996) for an interesting attempt to derive this difference between German and Japanese by a modification of the approach to Unambiguous Domination effects developed in section 7 below; Sauerland's basic strategy is to exploit the idea that the [scr] features of German and Japanese might differ with respect to interpretability, an assumption that I do not make here.

b. *[$_{\text{PP}}$ Über wen]$_1$ denkst du [$_{\text{NP}}$ ein Buch t_1]$_2$ sollte man mal t_2
about whom think you a book should one once
lesen ?
read

c. *Darüber$_1$ fragte sie sich [$_{\text{NP}}$ was für ein Buch t_1]$_2$ sie t_2 lesen sollte
about-that asked she REFL what for a book she read should

In (66-a) there is an illegitimate combination of *wh*-extraction from NP and NP scrambling; in (66-a) an illicit combination of *wh*-extraction from NP and NP topicalization occurs; finally, (66-b) illustrates the islandhood of *wh*-moved NPs for PP topicalization. In all these cases, the Unambiguous Domination requirement in (48) is fulfilled, because the positions occupied by XP_2 and XP_1 are different. But, as shown in some detail in chapter 4, the examples in (66) are ruled out because there is no well-formed derivation that respects cyclicity, locality, and economy constraints. Thus, a derivation of a sentence like, e.g., (66-a) in which *wh*-extraction of PP_1 follows scrambling of NP_2 violates the Barriers Condition; if the two movement operations apply in reverse order (i.e., *wh*-movement from NP precedes NP scrambling), the Strict Cycle Condition is not respected; and, finally, if (66-a) is the result of chain interleaving – extraction of PP_1 to a VP-adjoined position (i.e., scrambling), followed by scrambling of NP_2 to the pre-subject position, followed by *wh*-movement of PP_1 to SpecC –, the derivation is excluded because it involves a PUB effect (scrambling followed by *wh*-movement, applying to one and the same item), reducible to the interaction of Fewest Steps and Last Resort according to the approach developed in section 6 of chapter 2.

It seems that this analysis of Freezing effects straightforwardly extends to the examples in (62) – (65), and that the Unambiguous Domination requirement can therefore be assumed to hold for bound traces as well, but vacuously so. However, this conclusion might be premature. Consider an example like (62-a), where NP scrambling has taken place from a scrambled VP:

(67) *dass [$_{\text{NP}}$ den Aufsatz]$_1$ mal wieder [$_{\text{VP}}$ t_1 gelesen]$_2$ keiner t_2 hat
that the article$_{acc}$ once again read no-one has

If VP scrambling follows NP scrambling, the Strict Cycle Condition is violated. If, on the other hand, NP scrambling follows VP scrambling, the Barriers Condition is not fulfilled. However, there is a third derivation that does not seem to be excluded under the approach to Freezing developed in chapter 4, and this was in fact noted as a potential problem in subsection 6.2 of that chapter. This derivation involves chain interleaving as in (68) (= (118) from chapter 4):

(68) a. *D-Structure Representation*:
dass mal wieder keiner [$_{\text{VP}}$ [$_{\text{NP}}$ den Aufsatz]$_1$ gelesen]$_2$ hat

b. *NP_1 Scrambling to VP*:
dass mal wieder keiner [$_{\text{NP}}$ den Aufsatz]$_1$ [$_{\text{VP}}$ t_1 gelesen]$_2$ hat

c. *VP_2 Scrambling to IP*:
dass mal wieder [$_{VP}$ t_1 gelesen]$_2$ keiner [$_{NP}$ den Aufsatz]$_1$ t_2 hat

d. *NP_1 Scrambling to IP*:
dass [$_{NP}$ den Aufsatz]$_1$ mal wieder [$_{VP}$ t_1 gelesen]$_2$ keiner t'_1 t_2 hat

Here, NP_1 first undergoes local scrambling to VP_2; then, VP_2 is scrambled; and finally, NP_1 is scrambled to the position immediately following the complementizer. Since chain interleaving in this case involves two identical types of positions, it is permitted by Fewest Steps and Last Resort. In particular, intermediate adjunction of NP to VP in the first step is not excluded as a PUB effect, because this movement ends up in a position that is a potential checking position for the [scr] feature (this satisfies Last Resort) but does not actually result in checking of this feature (in accordance with the demands imposed by Fewest Steps). Thus, it seems that in order to rule out derivations like the one in (68), more must be said.[35] And more is indeed said by the Unambiguous Domination requirement – this condition is violated in the derivation in (68) (as well as in the derivation in (i) in the last footnote). Thus, if these considerations are on the right track, the principle of Unambiguous Domination even finds independent corroboration in the case of bound traces and not just with unbound traces.

5.2. Anaphors vs. Variables

5.2.1. Case-Driven Movement to SpecI

As it stands, (48) is a constraint on all traces, but so far I have discussed only examples involving variables, i.e., traces of A-bar movement. The question is: Is there evidence that Unambiguous Domination restricts the occurrence of traces of A-movement, i.e., anaphors? Note first that (48) does not preclude topicalization or *wh*-movement (indeed, any kind of A-bar movement) of a remnant XP when the antecedent of the unbound trace is located in an A-position. And indeed, as was noted in chapter 1 already, constructions of this type can be found in English (see Thiersch (1985, 13), Barss (1986,

---

[35] A similar point can be made with the derivation in (i):

(i) a. *D-Structure Representation*:
dass mal wieder keiner [$_{VP}$ [$_{NP}$ den Aufsatz]$_1$ gelesen]$_2$ hat

b. *VP_2 Scrambling to IP*:
dass mal wieder [$_{VP}$ [$_{NP}$ den Aufsatz]$_1$ gelesen]$_2$ keiner t_2 hat

c. *NP_1 Scrambling to VP*:
dass mal wieder [$_{VP}$ [$_{NP}$ den Aufsatz]$_1$ [$_{VP}$ t_1 gelesen]]$_2$ keiner t_2 hat

d. *NP_1 Scrambling to IP*:
dass [$_{NP}$ den Aufsatz]$_1$ mal wieder [$_{VP}$ t'_1 [$_{VP}$ t_1 gelesen]]$_2$ keiner t_2 hat

Here, NP scrambling takes place after VP scrambling, but since there is an additional intermediate operation of NP adjunction to VP, the Barriers Condition is not violated.

409), Browning (1987, 362ff), Saito (1989, 199), Höhle (1991, 170f), Huang (1993, 105ff), and others). Consider (69) (= (25)) first:

(69) [$_{\text{VP}}$ Criticized t_2 by his boss]$_1$ John$_2$ has never been t_1

Here, a remnant VP has undergone topicalization, stranding the subject NP *John* that had undergone Case-driven movement, and the result is well formed, as predicted by Unambiguous Domination. Under the Predicate-Internal Subject Hypothesis, the same conclusion of course applies with any kind of VP topicalization in English, as shown in (70) (= (24) from subsection 2.5 above):

(70) [$_{\text{VP}}$ t_1 Kicked the dog]$_2$ John$_1$ never has t_2

Next, consider an English example like (71), where *wh*-movement applies to a remnant AP, leaving behind the subject that has undergone Case-driven raising to SpecI (also, cf. footnote 10 in chapter 1 for putative counterexamples raised in Saito (1989) and Lasnik & Saito (1992)):

(71) [$_{\text{AP}}$ How likely t_1 to win the game]$_2$ is Mary$_1$ t_2 ?

Again, Unambiguous Domination does not rule out this combination of movement types, and again this result is empirically desirable.

Similarly, topicalization and *wh*-movement in German can apply to XPs in such a way that an unbound trace is created which has an antecedent in SpecI. This is shown for topicalization in (72-a) and for *wh*-movement in (72-b):[36]

(72) a. [Sich t_1 dumm vorkommen] sollte der Fritz$_1$ sicherlich nicht
REFL stupid strike-as should ART Fritz$_{nom}$ certainly not
b. [Wie t_1 dumm] ist sich der Fritz$_1$ vorgekommen ?
how stupid is REFL ART Fritz$_{nom}$ struck-as

Unbound traces in remnant XPs which have undergone scrambling are also not prohibited in German if they have an antecedent in an A-position (and not in a scrambling position; see above). This is shown in (73):

(73) ??dass [sich so richtig t_1 dumm vorgekommen] der Fritz$_1$ noch nicht ist
that REFL so really stupid struck-as ART Fritz$_{nom}$ not yet is

---

[36]See Sternefeld (1985; 1991a) for evidence that *vorkommen* (‘strike as,’ ‘appear’) is a raising verb in German. Note also that the direct *wh*-movement analogue of (72-a) is somewhat worse than (72-b):

(i) ?[Wie t_1 dumm vorgekommen] ist sich der Fritz$_1$?
how stupid struck-as is REFL ART Fritz$_{nom}$

I contend that the deviance of (i) has nothing to do with remnant movement but rather is due to the difficulty of understanding the bracketed items as a *wh*-phrase; i.e., it appears that (for some speakers) (i) involves illicit pied piping; cf. Webelhuth (1992, 116ff) and Lutz & Trissler (1992).

Examples like (73) are far from perfect; however, they have the intermediate status that arises in the case of predicate scrambling throughout, irrespective of the issue of remnant movement.

For the sake of completeness, let me point out that the same predictions as in English arise under the Predicate-Internal Subject Hypothesis. If subjects are base-generated VP-internally, then (74) must involve the generation of an unbound trace:

(74) [$_{\text{VP}}$ t$_1$ Geschlafen]$_2$ hat der Fritz$_1$ nicht t$_2$
slept has ART Fritz$_{nom}$ not

Remnant movement in this configuration meets the Unambiguous Domination requirement (48), since the landing site of the remnant VP and the subject NP differ. The same goes for cases of "complete" VP scrambling; cf. (26-b), repeated here as (75):

(75) ??dass [$_{\text{VP}}$ t$_2$ das Buch$_1$ gelesen]$_3$ keiner$_2$ t$_3$ hat
that the book read no-one has

If VP$_3$ contains an unbound subject trace in (75), the latter does not violate Unambiguous Domination because SpecI and an IP-adjoined position are formally different.

Thus, an occurrence of unbound anaphoric traces in remnant categories that have undergone topicalization or *wh*-movement is what we expect under (48), and this expectation is indeed fulfilled.[37]

Summarizing so far, condition (48) does not rule out A-bar movement of a remnant category if the trace included in the remnant has an antecedent in a SpecI position. This is a welcome result; however, we end up with exactly the same result if we assume that unbound anaphoric traces do not obey the Unambiguous Domination requirement in the first place. Thus, an interesting question is whether there is any positive evidence for (48) in the case of anaphoric traces. Relevant configurations would look as in (76),

---

[37] It has sometimes been argued, on the basis of the Predicate-Internal Subject Hypothesis, that subjects in German, unlike subjects in English, do not have to raise to SpecI overtly; see in particular Haider (1990; 1993). If this is so, it should be possible to topicalize a sequence consisting of, e.g., an intransitive verb and its subject NP. Given what has been said so far, it should even be possible to topicalize a transitive verb and its subject, leaving the object behind in a scrambling position. Relevant examples that illustrate these cases are (i-a) and (i-b) (both examples, together with their evaluation, are taken from Haider (1990, 97)):

(i) a. [$_{\text{VP}}$ Ein Außenseiter gewonnen]$_2$ hat da noch nie
an outsider$_{nom}$ won has there yet never

b. [$_{\text{VP}}$ Ein Außenseiter t$_1$ gewonnen]$_2$ hat das$_1$ da noch nie
an outsider$_{nom}$ won has that$_{acc}$ there yet never

Personally, I find these examples deviant (which would follow if subjects must undergo Case-driven movement at S-structure in German or if they are base-generated in a VP-external position in the first place), but it is at least worth noting that under the assumption that subjects are base-generated VP-internally and may stay in situ at S-structure, the judgements that Haider attributes to (i-ab) are compatible with Unambiguous Domination. See Haider (1990; 1993), Frey & Tappe (1991), Fanselow (1992), Truckenbrodt (1992), and literature cited in these works for further discussion of the issue at hand.

where a remnant XP γ has undergone movement to an A-position, e.g., SpecI, and also contains a trace with an antecedent in an A-position.

(76) *[$_{IP}$ [$_{\gamma}$ ... t_1 ...]$_2$ ... [$_{IP}$ YP_1 [$_{I'}$ ... t_2 ...

If the Unambiguous Domination requirement also holds for traces of A-movement, (76) is ruled out. But notice that (76) is independently ruled out, irrespective of whether XP is a remnant or not, since it involves a standard super-raising configuration. The conclusion to be drawn from this is that the case of remnant A-movement with an unbound A-trace is completely analogous to the case of remnant topicalization with an unbound topic trace discussed in subsection 4.4: The evidence from unbound traces of Case-driven movement is fully compatible with the Unambiguous Domination requirement, but this condition turns out not to be necessary in order to capture all the relevant data – topicalization of remnants and non-remnants (in fact, any kind of movement) across a topic is ruled out independently, and so is A-movement of remnants and non-remnants across a filled A-position.

5.2.2. Dative Movement

Consider now some data involving double object constructions in German for which I have not yet provided an explanation (cf. footnote 8):

(77) a. ??dass [$_{VP}$ (t'_1) ein Buch (t_1) gegeben]$_3$ die Claudia dem Peter$_1$ (nicht) t_3 hat
that a book$_{acc}$ given ART Claudia ART Peter$_{dat}$ not has

b. [$_{VP}$ (t'_1) Ein Buch (t_1) gegeben], das hat die Claudia dem Peter$_1$ (nicht) t
a book$_{acc}$ given that has ART Claudia ART Peter$_{dat}$ not

(77-a), in which a remnant VP which contains the direct object and V is scrambled, is much better than the corresponding example in which a remnant VP which contains the indirect object and V, or V alone, is scrambled (cf. the data in (6)); it is deviant only to the extent that predicate scrambling usually is. In (77-b) a remnant VP consisting of the direct object and V is left dislocated, and the result is fine, in contrast to left dislocation of the indirect object and V, or of V alone (cf. the examples in (19)). Thus, minimally contrasting pairs are (77-a) & (78-a) (= (6-b)) and (77-b) & (78-b) (= (19-b)):

(78) a. *dass [$_{VP}$ dem Peter t_2 gegeben]$_3$ die Claudia ein Buch$_2$ t_3 hat
that ART Peter$_{dat}$ given ART Claudia a book$_{acc}$ has

b. *[$_{VP}$ Dem Peter t_2 gegeben], das hat die Claudia ein Buch$_2$ t
ART Peter$_{dat}$ given that has ART Claudia$_{nom}$ a book$_{acc}$

This follows directly if we assume that in German indirect objects bearing dative Case are base-generated in the specifier position of a VP-shell that dominates the "real" VP, i.e., the one that is projected by the main verb. In (77-a) and (77-b), under this view, the smallest VP (containing only the direct object and the verb) has undergone movement, and a trace of the indirect object is not present. Such an approach is pursued by Truckenbrodt (1992). On the other hand, it is argued in Müller & Sternefeld (1994) and Müller (1995, ch. 4) that dative-bearing indirect objects in German have undergone (Case-driven) dative movement (dative shift) from a position asymmetrically c-commanded by the direct object to the specifier of a VP-shell. Under this assumption, an anaphoric trace t_1 is obligatorily present in the fronted VPs in (77). However, t_1 cannot violate the Unambiguous Domination requirement in (77-a) and (77-b), because its antecedent occupies the specifier position of the VP-shell, whereas the remnant VP is either scrambled (in (77-a)) or left-dislocated (in (77-b)). The question arises of whether or not t'_1 is also present in (77) (this trace would occupy the specifier of the VP-shell, i.e., the landing site of dative movement) – in other words, whether the VP-shell has been moved or only the lower VP. Indeed, if the VP-shell has been moved in (77), the antecedent of this trace (i.e., *dem Peter*) would have to occupy a scrambling position. But then, Unambiguous Domination would be violated. Hence, I conclude that what has undergone movement in (77) is not the VP-shell, with an unbound scrambling trace and a bound trace of dative movement, but only the lower VP, with merely an unbound trace of dative movement, as schematically depicted in (79) (where "DO" stands for direct object, and "IO" stands for indirect object):

(79) ... $[_{\text{VP}_2}$ DO t_1 V $]_3$ ... $[_{\text{VP}_1}$ IO_1 $[_{\text{VP}_2}$ Neg $[_{\text{VP}_2}$ t_3 $]$ $[_{\text{V}}$ – $]$ $]]$ ...

To sum up, the fact that the sentences in (77), unlike their counterparts in (78), do not violate Unambiguous Domination follows both under an analysis where indirect object NPs bearing dative Case are base-generated in a higher position than direct objects bearing accusative Case and under an analysis where indirect object NPs bearing dative Case have undergone dative movement.

5.3. X^0 Traces

Unambiguous Domination permits unbound X^0 traces to occur in moved remnant XPs; so, if nothing else is violated, such constructions are expected to be possible throughout. In what follows, I will briefly discuss a number of phenomena where the postulation of unbound head traces does indeed seem unavoidable.

5.3.1. VP Topicalization and Unbound V Traces in German

As observed by Fanselow (1991, 104; 1993, 66) and Lenerz (1995, 1267), German permits sentences like those in (80):[38]

(80) a. (Ich glaube) $[_{\text{TopP}}$ $[_{\text{PP}}$ mit dem Ball] $[_{\text{PP}}$ ins Gesicht] $[_{\text{Top'}}$ hat
I believe with the ball into the face has
er ihr geworfen]]
he$_{nom}$ her$_{dat}$ thrown

b. (Ich glaube) $[_{\text{TopP}}$ $[_{\text{NP}}$ Kindern] $[_{\text{NP}}$ Bonbons] $[_{\text{Top'}}$ gibt man
I believe children$_{dat}$ sweets$_{acc}$ gives one$_{nom}$
besser nicht]]
better not

c. $[_{\text{PP}}$ Im Mai] $[_{\text{NP}}$ jeden Tag] $[_{\text{NP}}$ der Anette] $[_{\text{NP}}$ einen Brief] $[_{\text{PP}}$
in May every day ART Anette$_{dat}$ a letter$_{acc}$
nach Hamburg] hätte er lieber nicht schicken sollen
to Hamburg had he$_{nom}$ better not send should

d. $[_{\text{NP}}$ Die Borussia] $[_{\text{PP}}$ in Führung] schoß dann Effenberg
the Borussia in the lead kicked then Effenberg

e. $[_{\text{Adv}}$ Vortrefflich] $[_{\text{NP}}$ die Figuren] führte der berühmte französische
superbly the pieces led the famous French
Diplomat Charles Maurice Talleyrand
diplomat Charles Maurice Talleyrand

f. $[_{\text{Adv}}$ Hilflos] $[_{\text{PP}}$ auf dem Platz] irrte Möller umher
helplessly on the field wandered Möller about

Since the prohibition against multiple topicalization in German is well established, Fanselow (1991; 1993) concludes that the fronted items in examples of this type must be part of one single constituent – a VP with an empty V head. Given this state of affairs, the most straightforward analysis of these examples looks as is shown for (80-b) in (81), where V-raising out of the VP (in this case, via I to C, in order to fulfill the V/2 requirement with topicalization in German) has taken place prior to VP topicalization:[39]

---

[38] I am grateful to Jochen Geilfuß (p.c.) for originally making me aware of this construction. Sentences like those in (80) do not seem to be accepted by all speakers of German. Haider (1990, 103; 1991, 28) and Sabel (1995, 140), e.g., find analogous examples ungrammatical. I have nothing to say about this apparent dialectal (or idiolectal) variation. Note incidentally that (80-d) is taken from an *SDR 1* football broadcast, March 1996; (80-e) can be found in Henscheid (1983, 144); and (80-f) is from a football commentary in the newspaper *Frankfurter Rundschau* (August 19, 1996).

[39] Cf. Fanselow (1991). It should be noted, though, that this is not actually Fanselow's (1993) analysis. In this follow-up work, he assumes that the fronted VP has an empty V head, which, however, is not a trace of the verb.

(81) [$_{\text{TopP}}$ [$_{\text{VP}_2}$ [$_{\text{NP}}$ Kindern] [$_{\text{NP}}$ Bonbons] t_1] [$_{\text{Top}}$ gibt$_1$] man besser nicht t_2]

The antecedent of the unbound verb trace t_1 is in a head position, but t_1 is not dominated by a category in a head position (VP is in a topic position). Hence, Unambiguous Domination does not rule out (81), and the wellformedness of the sentence is expected.[40]

A slight digression might be in order here. V-raising out of VP is clearly crucial in this approach to data like those in (80). However, whereas this assumption is uncontroversial (indeed, hardly avoidable) in the case of (81) (and (80-d), (80-e), and (80-f)), it is not quite as straightforward in the remaining two examples in (80); in these cases, it is a non-finite verb form (a participle, as in (80-a), or an infinitive, as in (80-c)) that must have undergone raising out of the VP. Here, I will not try to present a detailed analysis of these kinds of head movement (and of the apparent optionality that is involved); it may suffice to point out that these movement operations have indeed been suggested for independent reasons in the literature.[41] By the same reasoning, we are led to the conclusion that finite verbs in verb-final clauses at least have the option to move out of the VP at S-structure in German, given the wellformedness of examples like (82):

(82) [$_{\text{VP}_2}$ [$_{\text{NP}}$ Kindern] [$_{\text{NP}}$ Bonbons] t_1] denke ich [$_{\text{CP}}$ t'_2 dass man besser
children$_{dat}$ sweets think I that one$_{nom}$ better
nicht t_2 gibt$_1$]
not gives

Here, VP$_2$ has undergone topicalization as in (80-b), but this time, it is moved long-distance, and the finite verb is not in V/2 position but shows up at the end of the embedded clause. This shows that finite verbs must be able to leave the VP at S-structure in German even if they do not end up in V/2 position. If we make the further assumption that movement of finite verbs to a right-peripheral I node at S-structure is in fact obligatory in German (in contrast to raising of participles and infinitives), this

[40] Initially, one might think that the same phenomenon might show up with cases like (i), i.e., that NP$_1$ and NP$_2$ are not scrambled separately in front of the subject here, but rather that (i) instantiates VP scrambling, with the verb having left its own projection and undergone raising to I.

(i) dass [$_{\text{NP}}$ Kindern]$_1$ [$_{\text{NP}}$ Bonbons]$_2$ besser keiner gibt
that children$_{dat}$ sweets$_{acc}$ better no-one$_{nom}$ gives

However, it is unlikely that an analysis of (i) in terms of VP scrambling can be correct. The reason is that whereas nothing precludes multiple scrambling of NPs here (scrambling does not create an island and, being adjunction, can be iterated), (i) should have an intermediate status if it were derived by VP scrambling (recall the weak ban on predicate scrambling in German) and should only be possible with an I-topicalization intonation pattern (which is not the case).

[41] See, e.g., Sternefeld (1995) for arguments in support of the idea that participles can raise to a functional head Part in German.

might also provide a key to the solution of a well-known problem that is posed by the contrast in (83) (see Zwart (1993, 108) and Sabel (1995, 13-14), among others):

(83) a. [$_{\text{VP}}$ t_1 Gelesen]$_2$ denke ich [$_{\text{CP}}$ t'_2 dass das Buch$_1$ keiner t_2 hat]
read think I that the book$_{acc}$ no-one$_{nom}$ has

b. *[$_\alpha$ Las]$_2$ denke ich [$_{\text{CP}}$ t'_2 dass das Buch$_1$ keiner t_2]
read think I that the book$_{acc}$ no-one$_{nom}$

As shown in (83), remnant VP topicalization can apply long-distance with participles as heads (abstracting away from the possibility that VP in (83-a) might be reanalyzable as a higher functional projection; see the last footnote), but remnant topicalization may never apply with finite verbs as heads. Assuming obligatory V-to-I raising of finite verbs in German, there are only two possibilities as far as the nature of α in (83-b) is concerned: Either α is an X^0 category (a complex I node), in which case movement is barred by the prohibition against moving an X^0 category into an XP position (see subsection 2.1 of chapter 1), or α is an IP, with traces of the object and the subject. In the latter case, the structure of (83-b) is (84):

(84) *[$_{\text{IP}}$ t_3 t_1 Las]$_2$ denke ich [$_{\text{CP}}$ t'_2 dass [$_{\text{IP}}$ das Buch$_1$ [$_{\text{IP}}$ keiner$_3$ t_2]]]
read think I that the book$_{acc}$ no-one$_{nom}$

Here, NP_1 and NP_3 have undergone scrambling to IP prior to raising of the lowest IP segment. As such, this derivation does not violate Unambiguous Domination or any of the constraints discussed in chapter 4 (Barriers Condition, Strict Cycle Condition, etc.). However, it is an instance of IP topicalization, which is known to be systematically impossible in German, even with complete categories; cf.:

(85) *[$_{\text{IP}}$ Keiner das Buch las]$_1$ denke ich [$_{\text{CP}}$ t'_1 dass t_1]
no-one$_{nom}$ the book$_{acc}$ read think I that

Thus, whatever ultimately turns out to be responsible for the prohibition against IP topicalization in German will also cover ill-formed examples such as (84) – assuming that there is indeed obligatory overt V-to-I raising in German.

After this detour, let me now return to the main topic of this subsection, which is to accumulate evidence for unbound X^0 traces in remnant XPs.

5.3.2. VP Topicalization and Unbound V Traces in French

Essentially the same situation as in the German examples in (80) arises in French, as shown by Dekydtspotter (1992, 128). A relevant example is given in (86):

(86) ... et [$_{\text{VP}_2}$ t_1 sa tirade soigneusement] elle énonca$_1$ t_2
and her part carefully she recited

Given that only one constituent can undergo fronting to the pre-subject position in (86), we can again conclude that the fronted item must be a VP with a missing V head. French permits overt V-raising out of VP (see Emonds (1978), Pollock (1989), Chomsky (1991; 1993), and many others), so the existence of examples like (86) conforms to expectations under present assumptions: The unbound V trace t_1 in (86) does not violate Unambiguous Domination because it is not dominated by an X^0 category.

5.3.3. VP Movement and Unbound V Traces in English

If English differs from French in that a finite main verb does not leave its VP at S-structure (see the literature just cited), the prediction is that a construction that is analogous to (86) will not be found in English, independently of the issue of remnant movement and Unambiguous Domination. As shown by (87), this prediction is borne out:

(87) *[$_{VP_2}$ t_1 The article carefully] John read$_1$ t_2

(87) is ruled out not by Unambiguous Domination (which is fulfilled), but by whatever precludes overt V raising out of VP in English.[42]

That said, there is a construction in English where a finite main verb does indeed appear to have undergone raising out of the VP and into a pre-subject position, viz., the locative inversion construction. And in this case, remnant VP movement does indeed appear to be possible, as noted by Rochemont & Culicover (1991), who discuss examples like (88):

(88) [$_{VP_2}$ t_1 Into the room nude] walked$_1$ John t_2

The exact details of the analysis of (88) are disputable and not generally agreed on; but the postulation of two movement operations, one of which is remnant movement,

---

[42] A complication arises in the form of double object VPs, as in (i):

(i) *[$_{VP_2}$ t_1 John (t_V) a book] he did not give$_1$

At first sight, one might assume that (i) is amenable to the same analysis as (87). However, given Larson's (1988; 1990) approach to double object constructions, there is an additional VP_2-internal trace of the verb present in (i), viz., t_V. According to this view, VP_2 is a VP-shell, and t_V is the head of the lower VP. Thus, one might assume that what has undergone topicalization in (i) is not VP_2, but only the lower VP_3, as in (ii):

(ii) *[$_{VP_3}$ John (t_V) a book] he did not [$_{VP_2}$ give$_1$ t_3]

As it stands, (ii) violates neither the Unambiguous Domination condition (48) nor the requirement that a finite main verb in English must not leave VP (*give* is still within the VP-shell). Note, however, that the ungrammaticality of (ii) is a well-known problem in a Larsonian approach to double object constructions (see, e.g., Larson (1988; 1990) and Johnson (1991) for discussion); furthermore, it is strongly reminiscent of the problem posed by the immobility of small clause complements in English (see Williams (1983) and Stowell (1991), among many others), as in (iii):

(iii) *[$_{SC}$ John foolish]$_1$ I would never consider t_1

The general consensus seems to be that the illformednes of (ii) and (iii) is due to Case-theoretic reasons. For present purposes, it may suffice to leave it at that.

seems unavoidable. If this is so, the question arises of whether Unambiguous Domination licenses the unbound V trace created by remnant movement, and given that the landing sites of VP_2 and V_1 are formally different, it does, as required.

5.3.4. *VP Scrambling and Unbound V Traces in Japanese*

Koizumi (1995, 170) argues that sentences like (89) in Japanese cannot be derived by deletion operations like Gapping or Right Node Raising (assuming the latter to be an instance of leftward deletion; see Wilder (1996)), but must be derived by Across-the-Board raising of the verb out of VP:

(89) Mary-ga [$_{VP_1}$ John-ni ringo-o 2-tu t_4] to [$_{VP_2}$ Bob-ni banana-o 3-bon
Mary$_{nom}$ John$_{dat}$ apple$_{acc}$ 2-CL and Bob$_{dat}$ banana$_{acc}$ 3-Cl
t_4] ageta$_4$ (koto)
gave fact

Suppose that this is correct. Then, as Koizumi (1995, 173) observes, cases where the conjoined VPs (VP_1 and VP_2) undergo long-distance scrambling, thereby stranding the verb in the embedded clause, must be instances of remnant VP scrambling; cf.:[43]

(90) [$_{VP_\&}$ [$_{VP_1}$ John-ni ringo-o 2-tu t_4] to [$_{VP_2}$ Bob-ni banana-o 3-bon t_4]]
John$_{dat}$ apple$_{acc}$ 2-CL and Bob$_{dat}$ banana$_{acc}$ 3-Cl
Nancy-ga [$_{CP}$ Mary-ga $t_\&$ ageta$_4$ to] omotteiru (koto)
Nancy-$_{nom}$ Mary$_{nom}$ gave that believe fact

Clearly, the position of the antecedent of the unbound trace(s) t_4 in (90) and the position of the scrambled $VP_\&$ are not identical. Hence, Unambiguous Domination is respected, and the construction is permitted.

5.3.5. *Unbound N Traces in German*

The examples involving unbound X^0 traces that have been discussed so far have all involved V traces and remnant VP movement. In principle, we should expect to find analogous cases with unbound N traces. Indeed, there are constructions in German that involve topicalization of two internal arguments of N, with N itself remaining within IP. These constructions turn out to be amenable to exactly the same analysis as the data in (80). As a prerequisite for the analysis, let us briefly consider the internal structure of an NP with two internal arguments, such as (91-a):

(91) a. [$_{NP}$ ein Buch von Fritz über die Liebe]
a book by Fritz about the love

[43] I assume here that the two conjoined VPs form a larger VP constituent, perhaps with *to* ('and') as the head.

b. [$_{NP}$ ein [$_{N'}$ [$_{N'}$ Buch von Fritz] über die Liebe]]
c. [$_{NP}$ ein Buch$_1$ [$_{NP}$ von Fritz [$_{N'}$ t_1 über die Liebe]]]

On the basis of evidence from binding theory (as employed for VP-structure in Barss & Lasnik (1986) and Larson (1988)), Haider (1993, 23) argues that complex NPs of the type in (91-a) cannot be analyzed as in (91-b), but must have an NP-shell structure as in (91-c), with raising of the head noun *Buch* into the empty head of the NP-shell: Linear precedence corresponds to asymmetric c-command among the internal arguments in (91). If this is correct, the sequence *von Fritz über die Liebe* in (91) is an NP headed by the trace of the head noun. As such, it should in principle (i.e., if lexical and structural constraints on extraction from NP are met) be movable, thereby creating an unbound X^0 trace of category N. As shown by (92), this prediction is borne out:

(92) [$_{NP}$ Von Fritz [$_{N'}$ t_1 über die Liebe]]$_2$ hat er noch [$_{NP}$ kein Buch$_1$ t_2] gelesen
by Fritz about the love has he yet no book read

5.3.6. *Unbound N Traces in French*

Clitics in the Romance languages are another case in point. Consider the following example, which is due to Dekydtspotter (1992, 128):

(93) [$_{NP}$ Une photo t_1]$_2$ en$_1$ fut prise t_2
a photo of it was taken

Here, a combination of NP$_2$ raising to SpecI and cliticization of *en*$_1$ to I takes place, resulting in a grammatical structure. A derivation in which cliticization precedes A-movement of NP$_2$ does not violate any of the constraints assumed so far; in particular, it does not violate the Unambiguous Domination requirement because NP$_2$ and the clitic *en*$_1$ do not occupy the same type of position.[44]

5.3.7. *Unbound N Traces in Italian*

The same picture emerges with cliticization in Italian. As noted by Longobardi (1985), Cinque (1990), Rizzi (1990), and Grewendorf & Sabel (1994), a combination of remnant XP movement and cliticization out of XP is systematically available in this language. Consider, e.g., (94) (taken from Cinque (1990, 176)), where head movement of the object clitic precedes remnant VP topicalization:

---

[44] Note incidentally that a reverse combination of movement operations (i.e., A-movement preceding cliticization) would in fact not violate the Strict Cycle Condition as formulated in (52) in chapter 3, even though it involves lowering. The reason is that the only cyclic node created by lowering out of SpecX to X-Adj that is relevant for an item in an X-adjoined position is XP itself; however, X-Adj is part of the minimal residue of XP. Thus, the only constraint that would be violated in this derivation is the Barriers Condition.

(94) [$_{VP}$ Incontrato t_1]$_2$ non l_1'ha t_2
met (she) not him has

Similarly, head movement (cliticization) may precede remnant *wh*-movement, as in (95) (which is taken from Grewendorf & Sabel (1994, 303)):[45]

(95) [$_{AP}$ Quanto affezionata t_1]$_2$ gli_1 ela_3 ritenevate t_3 t_2 ?
how much affectionate him_{dat} her_{acc} (you) believe
'How affectionate to him do you believe her to be?'

Neither of these examples violates Unambiguous Domination, and given that other principles of grammar (especially locality, cyclicity, and economy constraints) are not violated either, their wellformedness is expected.

Of course, there are many more constructions where one might argue that an unbound X^0 trace is involved;[46] but the general picture should be clear by now: Unbound X^0 traces can never violate Unambiguous Domination if they are created by a remnant XP movement operation since the landing sites of XP movement and X^0 movement are always of different types. The question then is: Under which circumstances can X^0 traces actually violate Unambiguous Domination?

5.3.8. *X^0 Traces that Violate Unambiguous Domination*

Consider (96), a typical incorporation structure that may either be the result of overt head movement, i.e., adjunction to X^0 in the syntax (in which case t_Y is present), or the result of a base-generated complex head (in which case t_Y is absent); as before, linear order is not important in this schematic representation:

(96) ... [$_Z$ Y Z] ... (t_Y)

As noted by Baker (1988) and Chomsky (1991), among many others, it is generally not possible to move Y out of the complex head Z in such a configuration (but see Roberts (1991) for a different view). This follows directly from the Unambiguous Domination requirement. Suppose that Y moves in (96), as shown in (97):

---

[45]Under the Predicate-Internal Subject Hypothesis, AP in (95) actually has two unbound traces, viz., the indirect object clitic trace t_1 and a subject clitic trace t_3; similarly, VP in (94) has an unbound trace of *pro*.

[46]For instance, basically the same phenomenon as in French and Italian exists in Czech. As observed by Toman (1986, 134), both *wh*-movement and topicalization can apply to remnant NPs that contain clitic traces; cf.:

(i) a. [$_{NP}$ Kolik t_1]$_2$ $jich_1$ t_2 spi ve sklepě ?
how many of them sleep in cellar

b. [$_{NP}$ Pět t_1]$_2$ $jich_1$ dáme t_2 do sklepa
five of them we will put to cellar

The unbound traces t_1 here clearly fulfill Unambiguous Domination.

(97) *... Y ... [$_Z$ t'_Y Z] ... (t_Y)

In this case, t'_Y is an X^0 trace dominated by an X^0 category (i.e., by an item in an X^0 position), and ungrammaticality is expected to arise, as desired. Next, consider what happens if the lower Z segment undergoes head movement in the configuration in (96):

(98) *... Z ... [$_Z$ Y t_Z] ... (t_Y)

Here, t_Z is an X^0 trace dominated by an X^0 node (the higher Z segment), so under a strict reading of Unambiguous Domination, this principle is still violated, just as it is violated in (97), and ungrammaticality is predicted to arise. Again, I take this result to be justifiable. To sum up, then, it follows from Unambiguous Domination that excorporation is never possible.

This view yields some interesting consequences for constructions involving separable prefixes in German. These kinds of particles are characterized by the fact that they must show up adjacent to V (i.e., realized as a single phonological word) in V-final clauses (cf. (99-a)) but must be left behind by V/2 movement (cf. (99-b)):

(99) a. dass er die Tür zumacht
that he the door shut-makes

b. Dann macht$_1$ er die Tür zu t_1
then makes he the door shut

Suppose that separable prefixes are base-generated as heads of intransitive PPs (as assumed in chapter 1; cf. Emonds (1976) and Kayne (1985), and Grewendorf (1990) and Stechow (1992) for German). It then follows that *zu* ('shut') in (99-b) must be stranded in its base position (i.e., PP-internally) and cannot have undergone incorporation to V and be stranded there – in that case, V/2 movement applying to *macht* ('makes') would create a structure like (98) which is ruled out by Unambiguous Domination. As for V-final structures like (99-a), there are two possibilities. It could be the case that the complex phonological word *zumacht* is a result of optional overt P-incorporation into V; alternatively, complex word formation might indeed be a phonological process here, operating under P/V adjacency, and P stays in situ at S-structure.

Under these assumptions, there is a priori no reason why a separable particle (being a PP) should not be be able to undergo fronting alone. Similarly, we expect to find cases of remnant VP topicalization with an unbound V trace that differ from the examples in (80) that were discussed in subsection 5.3.1 above in that an NP and an intransitive PP (i.e., a particle) show up in the remnant VP. As observed by Uli Lutz (p.c.) (also see Stechow & Sternefeld (1988, 476) and Sabel (1995a, 15)), these constructions do indeed occur in German.[47]

---

[47] Many more examples can be found that instantiate the general form of (100-a) and (100-b); see, e.g., the data given in Paul (1919, 86) and Behagel (1932, 255). However, it should also be noted that many other verb-particle combinations do in fact not permit these constructions. Compare, e.g., the data in (i) (cf.

(100) a. [$_{PP}$ Zu]$_1$ hat er die Tür t_1 gemacht
shut has he the door$_{acc}$ made

b. [$_{VP}$ Die Tür zu t_1]$_2$ hat er t_2 gemacht$_1$
the door$_{acc}$ shut has he made

To sum up the evidence for Unambiguous Domination to be gained from data involving X^0 traces: Unbound V traces and N traces that are created by XP movement seem to occur in abundance across languages, which is expected by Unambiguous Domination. However, Unambiguous Domination rules out bound X^0 traces that are created by excorporation. Thus, given that this result is desirable, we can conclude that there is indeed positive evidence for Unambiguous Domination to be gained from the consideration of X^0 traces.

6. A Local Domain for Unambiguous Domination

Further amendments to Unambiguous Domination are not only possible but are indeed necessary upon closer inspection. So far, (48) does not mention a local domain in which variables must fulfill the Unambiguous Domination requirement. With this in mind, consider the bound traces in (101):

(101) a. [$_{NP}$ Wessen Frage [$_\gamma$ wen$_1$ du t_1 gesehen hast]]$_2$ hat t_2 dich
whose question whom you seen have has you
geärgert ?
annoyed

b. ?Wieder hat [$_\gamma$ dass den Fritz$_1$ alle t_1 mögen]$_2$ keiner t_2 bezweifelt
again has that ART Fritz$_{acc}$ all$_{nom}$ like no-one doubted

c. [$_{NP}$ Die Behauptung [$_\gamma$ keiner$_1$ habe t_1 davon gewußt]]$_2$ ist t_2
the claim no-one has$_{subj}$ of-that known is
natürlich falsch
of course wrong

In (101-a) t_1 has a *wh*-antecedent in a SpecC position and is dominated by a complex *wh*-NP which also occupies a SpecC position, thereby violating the Unambiguous Do-

---

Fanselow (1993, 69)); also, recall the discussion in subsection 2.4 of chapter 1.

(i) a. ?*[$_{PP}$ An]$_1$ sollte man die Anette lieber nicht mehr t_1 rufen
up should one ART Anette$_{acc}$ better not anymore call

b. *[$_{VP}$ [$_{NP}$ Die Anette an t_1]$_2$ sollte man lieber nicht mehr t_2 rufen$_1$
ART Anette$_{acc}$ up should one better not anymore call

Thus, the conclusion that lexical variation plays a role here seems unavoidable. What this variation is due to, however, is unclear to me (but it seems that the issue of semantic transparency of the verb-particle concatenation might be relevant – the semantics of the complex verb can be compositionally determined in (100), but not in (i)).

mination condition. Nevertheless, the sentence is well formed. Similarly, t_1 in (101-b) is a scrambling trace (with *den Fritz* as antecedent) and is dominated by a CP which has itself undergone scrambling. Again, the result is acceptable, contrary to what one would expect given (48). The same situation can be found in (101-c) – t_1 has an antecedent in a topic position and is dominated by a complex NP which has also undergone topicalization; still, ungrammaticality does not arise.

Basically the same kind of problem shows up with respect to unbound traces. Consider the following examples:

(102) a. ??[$_{\text{NP}}$ Welche Behauptung [$_\gamma$ t_1 zu lesen wisse sie nicht mehr
which claim to read knows$_{subj}$ she not anymore
was$_1$ sie versucht habe]]$_2$ ist t_2 falsch ?
what she tried has$_{subj}$ is wrong

b. Wieder hat [$_{\text{NP}}$ die Behauptung [$_\gamma$ t_1 zu lesen habe keiner das Buch$_1$
again has the claim$_{acc}$ to read has$_{subj}$ no-one the book
versucht]]$_2$ nur Peter t_2 geglaubt
tried only Peter$_{nom}$ believed

(102-a)

represents a case of an unbound trace of *wh*-movement which is ultimately dominated by a *wh*-phrase in SpecC. Nevertheless, this example is just as grammatical as, say, (55-a), which more or less corresponds to a non-embedded occurrence of γ in (102-a). Similarly, the unbound trace of scrambling in (102-b) may eventually be dominated by a category which has itself undergone scrambling, apparently in violation of (48).[48]

In order to reconcile the wellformedness of the examples in (101) and (102) with the assumptions so far, it looks as though a notion of locality must be introduced in (48), such that the traces in (101) and (102) are subject to the Unambiguous Domination requirement only within the local domain γ. To be concrete, I would like to suggest that the relevant domain be defined as follows:

(103) *Chain Domain*:
The chain domain of a trace t is the set of categories (irreflexively) included in the minimal node that dominates both t and its chain antecedent.

As can easily be verified, the chain domain of the variables t_1 in the examples in (101) and (102) is γ (or slightly smaller than γ) throughout. Within γ, however, these traces fulfill the Unambiguous Domination requirement. In conclusion, (48) should be modified as follows:

---

[48] As noted before, due to the strict nature of topic islands there are no unbound traces of topicalization. Thus, an example involving an unbound trace which would be analogous to (101-c) cannot be constructed.

(104) *Unambiguous Domination* (revised)
In its chain domain, an α-trace must not be α-dominated.

7. Deriving Unambiguous Domination

7.1. The A-Over-A Principle

Finally, let me turn to the third question posed at the end of section 4: Why does the Unambiguous Domination requirement hold? First, let us ask whether the Unambiguous Domination requirement in its representational formulation respects minimalist requirements or not. Arguably, it does: Although, as noted in section 4, its effects are exclusively S-structure ones, the copy theory of movement (that was adopted for independent reasons in chapter 3) makes it possible to assume that Unambiguous Domination holds at LF. Moreover, the formulation in (48) does not involve any non-basic, derived notions: The constraint talks about dominance, traces, and position types, and all these concepts are externally motivated.[49]

Still, certain aspects of Unambiguous Domination are reminiscent of other constraints that have been suggested over the years. This becomes particularly clear when we consider the derivational formulation of this principle that was given in (i) in footnote 27 and is repeated here in (105):

(105) *Unambiguous Domination* (derivational version):
In a structure ... [$_A$ ... B ...] ..., A and B may not undergo the same kind of movement.

Closer scrutiny reveals that (105) is basically a version of the A-over-A principle (see Chomsky (1964) and Ross (1967)), which can be formulated as follows:[50]

(106) *The A-Over-A Principle*:
If, in a structure ... [$_A$ ... B ...] ..., A and B belong to the same category, and a movement rule γ may affect that category, γ can only apply to A, not to B.

Thus, the crucial difference is that under the A-over-A principle A and B compete for movement if they belong to the same category type, whereas under Unambiguous Domination they compete for movement if they target the same kind of landing site. Now, as is well known since Ross (1967), the A-over-A principle cannot be maintained as

[49] A potential caveat is that the notion of chain domain that has turned out to be crucial (cf. the definition of Unambiguous Domination in (104)) might not meet the requirement of independent motivation; but see the discussion of domains in Chomsky (1993; 1995).

[50] This formuation is not Chomsky's original one. Rather, it is devised so as to make the similarity to (105) as clear as possible.

such because it is both too strong and too weak.[51] The conclusion I would like to draw from this is that the kinship of Unambiguous Domination and the A-over-A principle is suggestive in the sense that one may want to look for an independent reason for why Unambiguous Domination holds; but the A-over-A principle by itself cannot yet be this independent reason. Another candidate is (what is left of) the Shortest Path Condition.

7.2. The Shortest Path Condition, the Minimal Link Condition, Last Resort, and Superiority

Recall the transderivational Shortest Path Condition that was discussed in chapters 2 and 3 above (in subsections 3.3 and 6.2, respectively). The condition is repeated here (= (35) from chapter 2):

(107) *Shortest Path Condition*:
If two derivations D_1 and D_2 are in the same reference set and the movement paths of D_1 are shorter than the movement paths of D_2, then D_1 is to be preferred over D_2.

Thus far, I have discussed three applications of this condition, viz., (a) Yo-Yo movement in Ewe (cf. Collins (1994)), (b) Freezing effects with A-movement in English (cf. Chomsky (1995)), and (c) Superiority effects in English (cf. Chomsky (1993), Kitahara (1993), and others). As regards (a) and (b), I have argued that the Shortest Path Condition is superfluous because its work is done by the Strict Cycle Condition, which is independently motivated. As regards (c), I have concluded that due to its transderivational nature, the Shortest Path Condition is not compatible with the hypothesis that *wh*-in situ items obligatorily undergo raising to a SpecC$_{[+wh]}$ position at LF – a hypothesis that I maintain throughout this monograph. However, I have not presented an alternative account of the Superiority effect.

Chomsky (1995, 296-297) suggests dispensing with the transderivational Shortest Path Condition in favour of a simple derivational condition that nevertheless maintains the underlying concept. This condition is called Minimal Link Condition (MLC), and Chomsky assumes that it is part of the operation Move α itself; in contrast, I will assume that the MLC is incorporated into the Last Resort condition.

---

[51] The standard example given in Ross (1967/1986, 11-12) in order to show that the A-over-A principle is too strong is this:

(i) the reports [$_{NP_1}$ which] the government prescribes [$_{NP_2}$ the height of [$_{NP_3}$ the lettering on [$_{NP_4}$ the covers of t_1]]]

Here, NP_1 is affected by relativization even though it is embedded in NP_4, which is embedded in NP_3, which in turn is embedded in NP_2. Given the A-over-A principle, only NP_2 should be allowed to undergo relativization (which, of course, it can), and not any of the other NPs (which, however, also can).

The basic idea is that given a structure like that in (108), in which α has an unchecked feature of type F and β can check off α's F feature, raising of α to β is blocked if γ is closer to β and is also equipped with an unchecked F feature.[52]

(108) ... β ... [... γ ... [... α ...] ...] ...

Recall the revised version of Last Resort developed in section 6 of chapter 2 and assumed thereafter (= (134) from chapter 2):

(109) *Last Resort*:
α is raised to a position β only if β is a typical checking position for the lowest-ranked unchecked morphological feature of α.

By adding a version of Chomsky's (1995, 296-297; 311) derivational Minimal Link Condition (MLC), we obtain the following formulation of Last Resort:

(110) *Last Resort*:
α is raised to a position β only if (a) and (b) hold:
a. β is a typical checking position for the lowest-ranked unchecked morphological feature F of α.
b. There is no γ with an unchecked F feature that is closer to β.

For the time being, let us assume that "closeness" is understood in terms of length of paths (as defined in subsection 7.1 of chapter 3), such that γ is closer to β than α is if the path from γ to β is shorter than the path from α to β; I will return to this issue. Then, as can be easily verified, this formulation of the Minimal Link Condition in (110-b) does not say anything about Yo-Yo movement in Ewe or Freezing effects with A-movement. However, Superiority effects now follow from Last Resort.[53] To see this, consider again some typical examples, as in (111) and (112):

(111) a. I wonder [$_{CP}$ who$_1$ C [$_{IP}$ t$_1$ bought what$_2$]]
b. *I wonder [$_{CP}$ what$_2$ C [$_{IP}$ who$_1$ bought t$_2$]]

(112) a. Whom$_1$ did John persuade t$_1$ [$_{CP}$ to visit whom$_2$] ?
b. *Whom$_2$ did John persuade whom$_1$ [$_{CP}$ t$'_2$ to visit t$_2$] ?

Assuming that Case-driven movement has applied before *wh*-movement in all these examples, including the object movement cases (in apparent violation of Procrastinate; cf. Branigan (1992) and Chomsky (1995)), the (a)-requirement of Last Resort in

---

[52] This idea is identical in all relevant respects with the notion of feature-based relativized minimality that is developed and extensively argued for by Fanselow (1991, 76ff); also, compare Ferguson & Groat (1994).

[53] Compare, however, the cautious remarks in Chomsky (1995, 387: note 69).

(110) is respected throughout.[54] However, the (b)-requirement of Last Resort in (110) is violated by the derivations that create (111-b) and (112-b): In these cases, overt movement of the lower *wh*-XP$_2$ (*what*$_2$ and *whom*$_2$, respectively) to C$_{[+wh]}$ is prohibited because there is an intervening item that is closer to C$_{[+wh]}$ and that also has an unchecked [+wh] feature (*who*$_1$ and *whom*$_1$, respectively). Note that this MLC-based approach to Superiority phenomena is not susceptible to the criticism raised above against the approach in terms of a transderivational Shortest Path Condition. The reason is that the approach in terms of Last Resort in (110) is compatible with the assumption that *wh*-phrases undergo movement to SpecC$_{[+wh]}$ at LF. Consider, e.g., the LF of a sentence like (111-a), which is given in (113):

(113) I wonder [$_{\text{CP}}$ what$_2$ [$_{\text{CP}}$ who$_1$ C [$_{\text{IP}}$ t$_1$ bought t$_2$]]]

The *wh*-object *what*$_2$ can move across the intervening *wh*-subject *who*$_1$ to the minimal residue of C$_{[+wh]}$ (more specifically, to a CP-left adjoined position, given the assumptions in section 5 of chapter 4) without violating Last Resort, because although *who*$_1$ is closer to C$_{[+wh]}$ than *what*$_2$ is before *what*-movement takes place, the subject *wh*-phrase no longer has an unchecked [+wh] feature and for this reason does not block raising of the *wh*-object.

At this point, the question arises of what all this has to do with the main goal of this section, which is to try to derive the effects of Unambiguous Domination from independently motivated principles. I turn to this in the next subsection.

7.3. Unambiguous Domination and Last Resort

7.3.1. The Idea

It has been suggested by Kitahara (1994, 185 & 201-205), Fox (1995), and Koizumi (1995, 186-187) that the generalization in (47) (i.e., remnant XPs cannot undergo Y-movement if the antecedent of the unbound trace has also undergone Y-movement), and therefore (most of) the effects of the Unambiguous Domination requirement, can be derived from something like the MLC in (110-b).[55] In what follows, I would like

---

[54]Also, recall from subsection 3.4.1 of chapter 2 that I have assumed that the [+wh] features of *wh*-phrases are always weak – abstracting away from Bulgarian-type languages, it is the strength or weakness of C that creates cross-linguistic variation. As for English, I have suggested that the [+wh] feature of C, which is [+Interpretable] (i.e., does not disappear after checking), is strong to start with, but becomes weak after checking. If this is so, Procrastinate cannot serve to differentiate between legitimate and illegitimate *wh*-movement in (111) and (112).

[55]The actual implementation of this idea differs somewhat from author to author. Thus, Kitahara (1994, 185) proposes an economy constraint called "Restrict" that incorporates a statement akin to (110-b); Fox (1995) suggests that the generalization in (47) be derived by a transderivational Shortest Move condition; finally, Koizumi (1995, 187) formulates (110-b) as part of the operation Attract/Move, just like Chomsky (1995, 297).

to tentatively develop an approach to Unambiguous Domination along these lines. The approach builds on Kitahara's, Fox's, and Koizumi's basic insight and develops it further in a specific direction within the system of assumptions adopted so far.

First, recall the derivational version of Unambiguous Domination in (105), which is repeated once more:

(114) *Unambiguous Domination* (derivational version):
In a structure ... [$_A$ ... B ...] ..., A and B may not undergo the same kind of movement.

We can re-interpret the statement that "A and B may not undergo the same kind of movement" in this structure as "A and B may not check the same kind of feature." Furthermore, recall from section 6 that a local domain has proved necessary for Unambiguous Domination – illformedness results only when the target position of B is outside A (if B moves A-internally, this is compatible with Unambiguous Domination). In the structural environment recognized by (114), this means that the category that might check off the common feature of A and B must be outside A in order for Unambiguous Domination to become relevant at all; schematically, this is shown in (115) (with A replaced with γ, B replaced with α, and the category that checks off the feature F that triggers movement of both γ and α given as β, all for reasons of uniformity):

(115) ... β ... [$_\gamma$ ... α ...] ...

The crucial observation now is that in this context γ is invariably closer to β (in terms of length of path) than α is, the simple reason for this being that α is dominated by γ. Hence, we obtain exactly the same result as in the typical Superiority configuration in (116), where γ is closer to β than α is not because γ dominates α, but because γ asymmetrically c-commands α:[56]

(116) ... β ... [γ [... α ...] ...] ...

Given the revised Last Resort condition in (110), we expect that just as γ must move before α does in (116), γ-movement must precede α-movement in (115) if γ and α have the same unchecked feature F. Let us now see how the effects of Unambiguous Domination can be derived from Last Resort (in interaction with other constraints). I will begin by considering remnant scrambling.

7.3.2. Last Resort and Remnant Scrambling

Remnant scrambling must be systematically excluded if the antecedent of the unbound trace has also undergone scrambling, but not otherwise. A relevant pair of examples is

[56] As before, linear precedence by itself is not important in this configuration but is assumed to reflect asymmetric c-command relations.

repeated in (117) (cf. (28-a) and (29-b)):

(117) a. *dass [$_\alpha$ t_1 zu lesen]$_3$ keiner [das Buch]$_1$ t_3 versucht hat
that to read no-one the book$_{acc}$ tried has
b. dass [$_\alpha$ t_1 zu lesen]$_3$'s$_1$ keiner t_3 versucht hat
that to read it$_{acc}$ no-one tried has

The initial (i.e., pre-movement) structure of the illegitimate case of coherent infinitive scrambling in (117-a) looks roughly as in (118) (again, for the sake of simplicity, I abstract away from linear order and pretend that heads like I and V are left-peripheral in German):

(118) ... $X_{[scr]}$ ... $Y_{[scr]}$ ... [$_{\alpha_{[scr]}}$ ... $NP_{[scr]}$ ...] ...

Suppose that X = I, Y = V, X asymmetrically c-commands Y, and both heads can check off a [scr] feature in their respective minimal residues. By assumption, α (= the coherent infinitive in (117-a)) and NP (= *das Buch* in (117-a)) are equipped with a [scr] feature that must be checked off by overt movement. It now turns out that under Last Resort in (110) (in particular, under the MLC in (b)), there is no legitimate derivation left, given the theory of movement developed so far. The argument goes as follows: Clearly, the coherent infinitive α in (118) is closer to X than the NP included in α is; it is also closer to Y than NP is. Thus, any derivation in which NP moves first in (118) to either X or Y will violate Last Resort. Hence, we know that α must move first to either X or Y. Suppose α moves to X (= I), in accordance with Last Resort. Since the [scr] feature of α is now checked, Last Resort does not per se preclude scrambling of NP in the second step. The only remaining landing site for this movement is the minimal residue of Y (= V); but this is a position that is lower than the landing site of the preceding movement operation. Hence, what we have here is an illicit case of lowering in the syntax, excluded by the Strict Cycle Condition and the Barriers Condition (cf. chapters 3 & 4). Next, consider the reverse application of movement; i.e., α moves first again (as required by Last Resort), but this time it ends up in Y. In the second step, NP moves out of α to X, again in accordance with Last Resort. This derivation does not violate the Strict Cycle Condition. However, it violates the Barriers Condition. Recall from chapters 2 and 4 that XPs in derived positions are invariably barriers; at the point in this derivation at which NP extraction out of α applies, α is no longer in a complement position and is therefore a barrier that blocks subsequent movement. Thus, given the interaction of the revised Last Resort condition (incorporating the MLC), the Strict Cycle Condition, and the Barriers Condition, it follows that a structure like that in (118) cannot end up in a well-formed sentence (if NP moves out of α to either X or Y).

Consider now legitimate cases of remnant scrambling, as in (117-b). Here, the pre-movement structure is as in (119). The main difference to (118) is that NP does not have a [scr] feature to check off, but rather a "clitic" feature that one might abridge as

[cl].[57]

(119) ... $X_{[scr]}$ ... $Y_{[cl]}$ ... $[_{\alpha_{[scr]}}$ ... $NP_{[cl]}$ ...] ...

Here, NP is the closest category to the host Y that has an unchecked [cl] feature; α does not intervene simply because it does not have a [cl] feature (let alone an unchecked one). Hence, Last Resort does not require that α-movement precedes movement of NP (i.e., this time we are not forced into a violation of the Barriers Condition or the Strict Cycle Condition). So, NP moves first, cliticizing onto Y, and α moves second, adjoining to XP; and this does not violate any of the constraints adopted so far, as desired.

One might think that this approach may explain facts about remnant scrambling but runs into problems as soon as ordinary cases of multiple scrambling are taken into account, because it wrongly excludes multiple scrambling along the same lines as Superiority configurations with *wh*-movement. To see that this is not so, consider a typical case of multiple scrambling in German (= (95-a) from chapter 4):

(120) dass $[_{NP}$ dem Fritz $]_1$ $[_{NP}$ das Buch $]_2$ keiner $[_{VP}$ t_1 t_2 geben]
that ART Fritz_{dat} the book_{acc} no-one_{nom} give
wollte
wanted to

The structure before scrambling is as in (121); the relevant difference to (118) is that although the first NP asymmetrically c-commands the second NP, neither dominates the other.

(121) ... $X_{[scr]}$ ... $Y_{[scr]}$ ... $NP_{1[scr]}$ ... $NP_{2[scr]}$ ...

Since NP_1 (the dative NP) is closer to X and Y than NP_2 (the accusative NP) is, Last Resort requires that NP_1 moves first, before NP_2. But this has no negative consequences: Unlike what is the case with *wh*-movement in English, there is more than one potential landing site for scrambling in German. Furthermore, scrambling does not erect an island for another, subsequent scrambling operation. Therefore, NP_1 may first adjoin to YP, and NP_2 may then adjoin to XP in the second step (the reverse combination in which NP_1 adjoins to XP and NP_2 then adjoins to YP is precluded by the Strict Cycle Condition). Similarly, what is permitted under present assumptions is that a head like, e.g., Y in (121) may be able to check off more than one [scr] feature. In that case, both NP_1 and NP_2 can move to the minimal residue of Y, adjoining there in any order; the only way that Last Resort restricts these movements is by demanding

[57] I assume here that the [scr] feature and the [cl] feature are checked on different heads. However, in principle nothing depends on this – the conclusions would be the same in a language (or construction) where [scr] and [cl] are checked by one and the same head.

that NP_1 move first.[58]

To sum up so far, the revised Last Resort condition in (110) interacts with other constraints of the theory of movement adopted here in such a way that the restrictions on remnant scrambling fall out directly, apparently without negative consequences for other cases of scrambling (but see below). These results carry over immediately to remnant left dislocation (which I have analysed as left-adjunction to CP, i.e., as an operation that is formally indistinguishable from scrambling) – assuming, that is, that left dislocation is derived via movement of the left-dislocated item; cf. footnote 31. Let me next turn to remnant topicalization.

7.3.3. Last Resort and Remnant Topicalization

Consider first some well-formed cases of remnant topicalization, such as (122-a), where an unbound scrambling trace shows up (= (1-a)), and (122-b), where the unbound trace has a *wh*-antecedent in a SpecC position (= (55-a)):[59]

(122) a. [$_\alpha$ t_1 Zu lesen]$_3$ hat [$_{NP}$ das Buch]$_1$ keiner t_3 versucht
to read has the book no-one tried

b. ??[$_\alpha$ t_1 Zu lesen]$_2$ weiß ich nicht [$_{CP}$ was$_1$ sie t_2 versucht hat]
to read know I not what she tried has

The pre-movement structures look roughly as in (123-a) and (123-b), respectively:

(123) a. ... Top$_{[top]}$ ... Y$_{[scr]}$ ... [$_{\alpha_{[top]}}$ ... NP$_{[scr]}$ ...] ...
b. ... Top$_{[top]}$ ... C$_{[+wh]}$ ... [$_{\alpha_{[top]}}$ ... NP$_{[+wh]}$ ...] ...

As indicated, the features of the α-internal NP and of α itself are different. Hence, even though α is closer to Y/C than NP in both cases, Last Resort permits derivations where NP moves to Y/C first, and α-raising to Top takes place after that. Since this kind of derivation does not conflict with any other condition (e.g., the Barriers Condition, the Strict Cycle Condition, and the Fewest Steps condition are all respected), nothing precludes remnant topicalization in (122).

The case is different with sentences like (124) (= (45)):

(124) *[$_\alpha$ t_1 Zu lesen]$_2$ glaube ich [$_{TopP}$ [$_{NP}$ dieses Buch]$_1$ hat keiner t_2 versucht]
to read think I this book has no-one tried

Here, the pre-movement structure looks as follows:

[58] However, it seems that in order to derive an S-structure representation in which NP_1 is adjoined to X and NP_2 is adjoined to Y in (121), more must be said, e.g., that NP_1 with an unchecked [scr] feature that is in the minimal residue of Y does not count as closer to Y than NP_2 in situ anymore (cf. Chomsky (1995, 356)).

[59] As noted before, (122-b) is not actually fully grammatical, but its intermediate status is due to a Subjacency-like Barriers Condition violation and hence independent of remnant movement.

(125) ... $\text{Top}1_{[top]}$ ... $\text{Top}2_{[top]}$ ... $[_{\alpha_{[top]}}$ ... $\text{NP}_{[top]}$ ...] ...

Both α and the α-internal NP have an unchecked [top] feature. Last Resort as in (110) therefore requires that α move first. If α moves to Top_1, as in (124), subsequent topicalization of NP to Top_2 involves illicit lowering and thus creates a violation of the Strict Cycle Condition (and an additional violation of the Barriers Condition). If, on the other hand, α moves to Top_2 first (an option that is also permitted by Last Resort), subsequent topicalization of NP (which then must end up in the minimal residue of Top_1) violates the Barriers Condition: α is invariably a barrier at the point where extraction of NP applies. This explains why a sentence like (126) is impossible (also see (63)):

(126) *[$_{\text{NP}}$ Dieses Buch]$_1$ glaube ich [$_{\text{TopP}}$ [$_\alpha$ t_1 zu lesen]$_2$ hat keiner t_2 versucht]
this book think I to read has no-one tried

Thus, the approach to Unambiguous Domination effects in terms of the revised Last Resort condition in (110) makes correct predictions for remnant topicalization. However, as was already noted above, the illformedness of examples like (124) and (126) would follow independently of the issue of remnant movement. Thus, due to the strict nature of topic islands in German, (124) would be impossible even if α were to undergo some other kind of movement. Similarly, (126) would come out as ungrammatical if NP_1 were affected by some other movement type, e.g., *wh*-movement. In other words, although the revised Last Resort condition does not make wrong predictions as far as remnant topicalization options are concerned, it seems that we do not really need it in this domain – assuming the earlier version of Last Resort (cf. (109)), the illformedness of (124) and (126) would still follow (from whatever accounts for the fact that there is no movement across a filled SpecTop position in German, and from the fact that bound traces are subject to Freezing effects, as shown in chapter 4).

7.3.4. Last Resort and Remnant Wh-Movement

Finally, let me turn to *wh*-movement. A well-formed case of remnant *wh*-movement in German is repeated in (127) (= (15-a)):

(127) [$_{\text{NP}}$ Was für ein Buch t_1]$_2$ hast du [$_{\text{PP}}$ über die Liebe]$_1$ t_2 gelesen ?
what for a book have you about the love read

The pre-movement structure is this:

(128) ... $\text{C}_{[+wh]}$ ... $\text{Y}_{[scr]}$ ... [$_{\text{NP}_{[+wh]}}$ ... $\text{PP}_{[scr]}$ ...] ...

NP is closer to both C and Y than PP is, but the unchecked movement features of NP and PP are different ([scr] vs. [+wh]). Therefore, Last Resort does not force us to move NP first (which would invariably result in a Barriers Condition violation and also in a Strict Cycle Condition violation); rather, PP may move to Y first, thereby checking its [scr] feature, and NP may then move to C, checking the [+wh] feature, in accordance with the Barriers Condition, strict cyclicity, etc.[60]

The canonical ill-formed cases of remnant *wh*-movement in English and German are repeated in (129-a) and (129-b) (= (35-a) and (38)):

(129) a. *[$_{\text{NP}}$ Which book about t_1]$_2$ don't you know [$_{\text{CP}}$ who$_1$ to read t_2] ?
b. *[$_{\text{NP}}$ Was für ein Buch t_1]$_2$ fragst du dich [$_{\text{CP}}$ [$_{\text{PP}}$ über wen]$_1$ du t_2
what for a book ask you REFL about whom you
lesen sollst] ?
read should

The initial structure before *wh*-movement has the following form:

(130) ... $C1_{[+wh]}$ ... $C2_{[+wh]}$ ... [$_{\text{NP}_{[+wh]}}$ ... $\text{XP}_{[+wh]}$ ...] ...

Now, Last Resort in (110) demands that NP moves first – NP is closer to both C nodes than XP is and has the same unchecked movement feature. If NP moves to C_1, as in the examples in (129), subsequent XP extraction is lowering and is precluded by the Strict Cycle Condition and the Barriers Condition. Thus, there is no derivation of these sentences that respects both Last Resort and the combined Strict Cycle Condition/Barriers Condition. On the other hand, suppose that NP in (130) moves not to C_1 but to C_2, followed by XP extraction to C_1. Again, the Last Resort condition forces a violation of the Barriers Condition in this case (though not of the Strict Cycle Condition), and the result is expected to be ill formed. Relevant examples are given in (131-a) for English and in (131-b) (= (64-a)) for German:[61]

(131) a. ?*Who$_1$ don't you know [$_{\text{CP}}$ [$_{\text{NP}}$ which book about t_1]$_2$ to read t_2] ?
b. *Worüber$_1$ hast du gesagt [$_{\text{CP}}$ [$_{\text{NP}}$ was für Bücher t_1]$_2$ er t_2 gelesen
about what have you said what for books he read
hat] ?
has

Thus, on the basis of (130), Last Resort forces a derivation of a sentence like (131-a) in which *wh*-movement of NP precedes *wh*-movement of XP (= *who*$_1$), in violation of the Barriers Condition. However, if the two movement operations were to apply in reverse order (i.e., *who*$_1$ first raises to the matrix SpecC_1 position, and NP_2 subsequently moves to the embedded SpecC_2 position), we would independently predict a violation of the Strict Cycle Condition. This means that as with the topicalization

[60] Similar considerations apply in the case of combinations of remnant *wh*-movement and extraposition in English, as discussed in subsection 3.3.4.

[61] A remark concerning the status of (131-a) in English is due. It is sometimes assumed that examples that are structurally analogous to the one at hand are better than, e.g., (129-a). If so, this could not be accounted for under present assumptions. However, since the contrast is at best a very subtle one, and since no such contrast shows up in German, I think that the approach developed here is not directly threatened by these judgements of sentences like (131-a).

example in (126), the revised Last Resort is not actually necessary in order to rule out the sentences in (131)– these examples are excluded as involving Freezing effects along the lines of what has been said in chapter 4, irrespective of whether both moved items undergo the same movement operation or not.

To sum up so far, the revised Last Resort condition in (110) that integrates an indepedently motivated derivational shortest path requirement, the MLC (cf. the discussion of Superiority effects in English), correctly predicts the restrictions on remnant movement that are the main topic of this chapter. It does so by allowing only derivations that violate other principles, viz., the Barriers Condition and the Strict Cycle Condition. Thus, recall from chapter 4 that remnant movement may "normally" trigger Anti-Freezing effects; now, given the approach developed here in terms of the revised Last Resort condition, it turns out that remnant movement may indeed sometimes have to induce a Freezing effect – this is so if the remnant XP and the item that is included in it have the same kind of unchecked movement feature. As with the Unambiguous Domination requirement, it turns out that there are a few redundancies with bound traces, in the sense that constructions that are ruled out by the (b)-requirement of Last Resort in (110) would be ruled out anyway, e.g., by the Strict Cycle Condition; but I do not believe that these redundancies are particularly troublesome.

More generally, then, if this approach is viable, it looks as though we can dispense with Unambiguous Domination. However, it should be noted that there are a number of problems that have to be solved before the enterprise of deriving Unambiguous Domination from Last Resort can be viewed as fully successful. I turn to this issue in the following subsection, noting the problems but not attempting to present full-fledged solutions. In addition, I also address a possible further modification of Last Resort that allows us to derive the constraint on bi-directional adjunction (94) of chapter 4.

7.4. Problems and Consequences

7.4.1. Locality/Last Resort Interactions

The first question we may ask ourselves is how the revised Last Resort condition in (110) that integrates the MLC interacts with the theory of locality adopted so far. The revised Last Resort condition is repeated in (132):

(132) *Last Resort*:
α is raised to a position β only if (a) and (b) hold:
a. β is a typical checking position for the lowest-ranked unchecked morphological feature F of α.
b. There is no γ with an unchecked F feature that is closer to β.

The Barriers Condition is violated in the examples in (133) which involve a *wh*-island violation with an argumental *wh*-phrase and which have an intermediate (Subjacency-like) status ((133-a) = (33-b) from chapter 2):

(133) a. ??Who$_3$ do you wonder [$_{CP_1}$ whether John said [$_{CP_2}$ t'_3 C t_3 solved the problem]] ?
b. ?*Which book$_3$ do you wonder [$_{CP_1}$ [$_{NP}$ who]$_4$ t_4 said [$_{CP_2}$ t'_3 that Mary read t_3]] ?

The first thing to note is that we do not want these examples to be ruled out by Last Resort for a very simple reason: Locality violations that arise in the derivation lead to weak ungrammaticality, and only locality violations that persist at LF result in full ungrammaticality. Given that the moved item is an argument that has established an "offending" intermediate trace, the intermediate status of (133-a) and (133-b) is derived under the Barriers Condition (cf. subsection 3.2.5 of chapter 2). However, in order to account for the strong degree of ungrammaticality that arises with Unambiguous Domination effects, it is a prerequisite that violations of Last Resort are always strong.[62] So, Last Resort must not rule out the sentences in (133). Fortunately, this result does indeed follow.

Thus, suppose that what happens in (133-a) is this: First, NP$_3$ moves to SpecC$_2$ (without actually checking the [+wh] feature); since the presence of *whether* blocks the use of SpecC$_1$ as an escape hatch (see Chomsky (1986)), NP$_3$ moves directly to the matrix clause SpecC position in the next step, where it checks off its [+wh] feature, in violation of the Barriers Condition. Note that neither of these movement operations violates Last Resort. The reason is that both movement operations end up in typical checking positions for an unchecked [+wh] feature and, more importantly in the present context, that in each case of *wh*-movement, the moved *wh*-item is closer to the target C position than any other *wh*-item with an unchecked [+wh] feature is (*whether* does not have an unchecked [+wh] feature).

A similar result is obtained in the case of (133-b). The pre-movement structure of this sentence looks as in (134):

(134) $C_{0[+wh]}$ ... $C_{1[+wh]}$ ... [$_{NP}$ who]$_{4[+wh]}$ ... C_2 ... [$_{NP}$ which book]$_{3[+wh]}$

Again, NP$_3$ moves to the typical [+wh] feature checking position SpecC$_2$ first, but does not yet check off its [+wh] feature (since CP$_2$ is a declarative clause). Second, NP$_4$ undergoes local movement to the specifier of C$_1$, thereby checking off its [+wh] feature. And finally, NP$_3$ moves from the SpecC$_2$ position to the matrix SpecC$_0$ position, checking off its [+wh] feature there. This derivation does not violate Last Resort since NP$_4$ does not count as an intervening element (for the purposes of the MLC) anymore after its [+wh] feature has been checked.[63]

---

[62] The same conclusion can be drawn on the basis of (strong) PUB effects as they are discussed in section 6 of chapter 2; recall that in that case, too, the Last Resort condition was relevant.

[63] At least, this is so if we ensure that NP$_4$ does not count as closer to C$_2$ than NP$_3$ in (134) – more generally, it must be guaranteed that only XP$_2$, and never XP$_1$, can undergo movement to the minimal residue of a head Z in a structure like (i), even if the path from XP$_1$ to Z is shorter than the path from XP$_2$

Nevertheless, these considerations lead to an undesirable prediction of Last Resort for sentences like those in (135) in English:

(135) a. ??Which car_1 did John tell you [$_{CP}$ how_2 (t'_1) to fix t_1 t_2] ?
b. *How_2 did John tell you [$_{CP}$ which car_1 to fix t_1 t_2] ?

The pre-movement structure of both these examples is depicted schematically in (136):

(136) ... $C_{1[+wh]}$ ... $C_{2[+wh]}$ ... *which car*$_{[+wh]}$ ... *how*$_{[+wh]}$ ... ...

Suppose that *how* is adjoined to VP, i.e., base-generated in a higher position than *which car*. In that case, it can undergo successive-cyclic movement to the minimal residue of C_1 without violating Last Resort; in the next step, *which car* moves to SpecC_2, again satisfying Last Resort but violating the Strict Cycle Condition. Hence, the resulting S-structure representation, viz., (135-b), is impossible. In this case, the effect of Last Resort is unproblematic, since the sentence is indeed completely ungrammatical.[64] Suppose next that *how* moves to the minimal residue of C_2 in the first step, again an option permitted by Last Resort. In that case, subsequent long-distance movement of *which car* violates the Barriers Condition; the result is (135-a).

However, a problem is now emerging: The offending trace created by this raising cannot be t_1 – given Chomsky & Lasnik's (1993) account of argument/adjunct asymmetries in terms of uniform chains and Fewest Steps (that was assumed above in chapter 2), an intermediate trace t'_1 that induces the locality violation and can (in fact, must) itself be deleted on the way to LF is needed in (135-a). I have not committed myself to any specific hypothesis about what kind of position *which car* raises to before it undergoes *wh*-movement, but it seems clear that this position is higher than the VP-adjoined position initially occupied by *how* (a possibility envisaged above was that this position is SpecTop, available as an intermediate landing site of *wh*-movement in infinitives). Under this assumption, it must actually be *which car*, not *how*, that is *wh*-moved first,

to Z:

(i) ... XP_1 [... Z ... [... XP_2 ...] ...] ...

In other words, Last Resort must not accidentally force illicit lowering that is then excluded as a violation of the Strict Cycle Condition. To accomplish this result, we have to revise the (b)-requirement of (132) as in (ii):

(ii) There is no γ with an unchecked F feature that is closer to and c-commanded by β.

For reasons of simplicity, I will abstract from this modification in what follows, but I will tacitly presuppose it.

[64] But of course, the revised Last Resort condition is not necessary to achieve this result; a reverse combination of movement operations in (135-b) would have involved a violation of locality requirements, with the SpecC_2 position blocked and strong ungrammaticality arising because long-distance adjunct movement is involved.

because of Last Resort. As before, two options arise. On the one hand, *which car* may move to the embedded SpecC$_2$ position, and *how* subsequently raises to the matrix SpecC$_1$ position, thereby creating the S-structure representation in (135-b). This sentence is strongly ill formed, and this follows because of a Barriers Condition violation with adjunct movement. On the other hand, however, *which car* may also move to the matrix SpecC$_1$ position first, followed by movement of *how* to the embedded SpecC$_2$ position in the second step. This derivation creates (135-a) which, as we have seen, has an intermediate status. Unfortunately, though, the best derivation that generates it (i.e., a derivation that respects Last Resort) still violates the Strict Cycle Condition, and therefore we would expect (135-a) to be completely ungrammatical, contrary to fact. In sum, it seems that in order to solve this problem, we have to make sure that movement of *how* to SpecC$_2$ may precede movement of *which car* to SpecC$_1$. This can be accomplished in one of two ways: Either the revised Last Resort condition in (132) has to be given up, or the notion of closeness must be modified in such a way that *which car* does not count as closer to C$_1$ and C$_2$ than *how* is even though it is located in a higher position (at the relevant stage in the derivation).

The same problem shows up independently of considerations related to argument/adjunct asymmetries with sentences like (137):

(137) ??Whom$_1$ don't you know [$_{\text{CP}_3}$ how$_2$ to tell t$_1$ [$_{\text{CP}_4}$ t$'_2$ to fix the car t$_2$]] ?

The initial structure of this sentence before any *wh*-movement operation looks as in (138):

(138) ... C$_{0[+wh]}$ ... [C$_{3[+wh]}$ ... whom$_1$ ... [C$_4$ ... how$_2$ ...] ...] ...

The adjunct *how*$_2$ is generated in CP$_4$, below *whom*$_1$, which is generated in CP$_3$. First, *how*$_2$ moves to C$_4$. But then, Last Resort demands that *whom*$_1$ moves next, and *how*$_2$ can move only afterwards. To create the string in (137), movement of *whom*$_1$ must end up in the minimal residue of C$_0$; but then, subsequent raising of *how*$_2$ to SpecC$_3$ violates strict cyclicity. Therefore, we incorrectly predict strong illformedness in (137), due to either a Strict Cycle Condition violation or a Last Resort violation, and cannot derive its actual intermediate status, which would ideally be attributed to violations of the Barriers Condition: The derivation in which only locality constraints are violated also violates Last Resort, under present assumptions (in this derivation, *how*$_2$ first moves to the embedded SpecC$_3$ position, followed by long-distance raising of *whom*$_1$ to SpecC$_2$). Again, the conclusion is that either Last Resort is wrong (in the form it takes in (132)) or the notion of closeness is not simply definable in terms of length of paths.

Naturally, in order to maintain the Last Resort-based account of Unambiguous Domination effects with remnant movement, the latter solution (i.e., modifying the notion of closeness) will be the preferred one; but I am not trying to advance a specific proposal here. Rather, let me now turn to another (albeit related) problem that shows up with multiple questions in German.

7.4.2. Superiority in German

German normally does not exhibit Superiority effects, in striking contrast to English (recall (111)). This is shown for co-arguments (objects and subjects) by the data in (139):[65]

(139) a. (Ich weiß nicht) [$_{CP}$ was$_1$ wer$_2$ t$_1$ behauptet hat]
I know not what$_{acc}$ who$_{nom}$ claimed has
b. (Ich weiß nicht) [$_{CP}$ wer$_1$ t$_1$ was$_2$ behauptet hat]
I know not who$_{nom}$ what$_{acc}$ claimed has

Similarly, examples with two *wh*-objects that correspond to the English data in (112) (where one object originates in the embedded infinitive, and the other one in the matrix clause) are immune to Superiority effects – either object may undergo S-structure *wh*-movement:[66]

(140) a. Was$_1$ hast du wen$_2$ überzeugt [$_{CP}$ t$'_1$ PRO t$_1$ zu kaufen] ?
what$_{acc}$ have you whom$_{acc}$ convinced to buy
b. Wen$_2$ hast du t$_2$ überzeugt [$_{CP}$ PRO was$_1$ zu kaufen] ?
whom$_{acc}$ have you convinced what$_{acc}$ to buy

These data are obviously problematic for any account that relates Superiority effects in languages like English to the MLC. The problem becomes even more pressing in the specific approach currently under consideration. The MLC is now part of the revised Last Resort condition, and it does not seem possible to maintain the view that this condition for some reason simply does not hold in German – after all, restrictions on remnant movement in German that were earlier derived by Unambiguous Domination now follow under the MLC, integrated into Last Resort. What is more, given Unambiguous Domination effects with *wh*-movement in German (cf. (129-b), repeated here as (141)), we seem to have direct evidence that *wh*-movement in German is restricted by Last Resort in the formulation given in (132):

(141) *[$_{NP}$ Was für ein Buch t$_1$]$_2$ fragst du dich [$_{CP}$ [$_{PP}$ über wen]$_1$ du t$_2$
what for a book ask you REFL about whom you
lesen sollst] ?
read should

To confuse things even more, it turns out that German does indeed exhibit Superiority effects if the two *wh*-phrases in question are base-generated in different finite clauses,

[65] See Haider (1983; 1986; 1993), Grewendorf (1988), and Bayer (1990; 1990a), among many others.

[66] See Fanselow (1991, 330) and Müller (1995, 323).

as in (142):[67]

(142) a. Wer$_1$ hat t$_1$ geglaubt [$_{CP}$ dass der Fritz wen$_2$ mag] ?
who$_{nom}$ has believed that ART Fritz$_{nom}$ whom$_{acc}$ likes
b.?*Wen$_2$ hat wer$_1$ geglaubt [$_{CP}$ t$'_2$ dass der Fritz t$_2$ mag] ?
whom$_{acc}$ has who$_{nom}$ believed that ART Fritz$_{nom}$ likes

As a final piece of problematic evidence in this context, consider a sentence like (143):

(143) [$_{PP}$ Über wen]$_1$ hast du [$_{NP}$ was für ein Buch t$_1$]$_2$ gelesen] ?
about whom have you what for a book read

Here, a *wh*-PP embedded in a *wh*-NP (that is transparent for extraction, due to indefiniteness and the right choice of the governing verb) undergoes *wh*-movement, yielding a well-formed sentence, but in apparent violation of the Last Resort condition (note that the case is completely parallel to (141) under present assumptions).[68]

Thus, we have conflicting evidence as regards the validity of the Last Resort condition in (132) in German: Data such as those in (139), (140), and (143) seem to suggest that MLC effects do not show up with *wh*-movement in German, whereas data such as (141) and (142-b) seem to suggest that they do. Note that the distinction is not one between Superiority-type constructions (as schematically depicted in (144-a)) and remnant movement-type constructions (as in (144-b)) – both kinds of constructions may or may not lead to well-formed sentences in German, as we have just seen.

(144) a. $C_{1[+wh]}$ ... $XP_{3[+wh]}$ ... $XP_{4[+wh]}$ ...
b. $C_{1[+wh]}$ ... $C_{2[+wh]}$ ... [$_{XP_{3[+wh]}}$ ... $XP_{4[+wh]}$ ...] ...

In view of this, it seems to me that again the best way to go is to try to relate the apparent (ir-) relevance of Last Resort to the notion of closeness, in such a way that Last Resort blocks the ill-formed examples (141) and (142-b), but not the well-formed examples in (139), (140), and (143). Given that the ill-formed *wh*-movement operations involve long-distance steps across finite clause boundaries, and the well-formed *wh*-movement operations typically employ only (relatively) local movements, this enterprise does not strike me as impossible (but I will not undertake it here, and closer

---

[67]The observation is, I believe, due to Werner Frey (p.c.); also see Büring & Hartmann (1994, 62) and Müller (1995, 324). Note also that (142-b) is well formed if the *wh*-phrase *wer*$_1$ is replaced by a [–wh] expression such as *Karl*; so, it must be a genuine Superiority effect that is at work here, and not some other factor that might bar extraction.

[68]Incidentally, the wellformedness of (143) at first sight also poses a problem for the idea that Unambiguous Domination applies at LF (cf. section 4 above). The reason is that the *wh*-trace t$_1$ will end up dominated by a category in a *wh*-position at LF (after *wh*-raising of NP$_2$). In contrast, (143) would be unproblematic if we were to assume that Unambiguous Domination applies at S-structure. However, this problem for an LF-based Unambiguous Domination condition disappears if we take into account that the [+wh] position of PP$_1$ (SpecC) and the [+wh] position of NP$_2$ (CP-Adj) are formally different.

scrutiny shows that it is not as straightforward as one would hope). Furthermore, it is evident that the definition of closeness must be parametrizable – otherwise, cross-linguistic variation with respect to Superiority phenomena with clause-mates would have to remain a mystery.[69]

That said, let me turn to a third and final problem raised by the present approach based on Last Resort.

7.4.3. Bound Scrambling Traces

The following example has already been discussed in some detail in subsection 5.1 of this chapter (cf. (67); also see subsection 6.2 of chapter 4):

(145) *dass [$_{\text{NP}}$ den Aufsatz]$_1$ mal wieder [$_{\text{VP}}$ t$_1$ gelesen]$_2$ keiner t$_2$ hat
that the article$_{acc}$ once again read no-one has

I have argued that a derivation as in (146) must be excluded and that this derivation does not violate any of the locality, economy, and cyclicity constraints recognized in chapters 2 – 4.

(146) a. *D-Structure Representation*:
dass mal wieder keiner [$_{\text{VP}}$ [$_{\text{NP}}$ den Aufsatz]$_1$ gelesen]$_2$ hat
b. *NP$_1$ Scrambling to VP*:
dass mal wieder keiner [$_{\text{NP}}$ den Aufsatz]$_1$ [$_{\text{VP}}$ t$_1$ gelesen]$_2$ hat
c. *VP$_2$ Scrambling to IP*:
dass mal wieder [$_{\text{VP}}$ t$_1$ gelesen]$_2$ keiner [$_{\text{NP}}$ den Aufsatz]$_1$ t$_2$ hat
d. *NP$_1$ Scrambling to IP*:
dass [$_{\text{NP}}$ den Aufsatz]$_1$ mal wieder [$_{\text{VP}}$ t$_1$ gelesen]$_2$ keiner t$'_1$ t$_2$ hat

However – so the argument in subsection 5.1 goes – the resulting representation in (146-b) violates Unambiguous Domination due to illicit domination of the scrambling trace t$_1$ by a scrambled item (VP$_2$). The question now is: Does the ungrammaticality of sentences like (145) that exhibit bound scrambling traces in scrambled items still follow under the present approach in terms of Last Resort? The crucial operation is

[69]Clearly, defining the notion of closeness (for the purposes of the MLC/Last Resort) exclusively via the concept of length of paths (or minimal c-command) is theoretically more attractive than resorting to auxiliary, non-basic assumptions whose task is to render the notion of closeness more “liberal” by creating equidistance of two items that stand in an asymmetric c-command relation. However, it is at least worth bearing in mind that the problem is far from being confined to the case at hand. For instance, the consideration of multiple clause-internal A-movement operations has often been viewed as necessitating a concept of “equidistance” or “minimal domain” in the MLC; cf., among others, Chomsky (1995, 356-358) (who points out that a strict notion of closeness that does without minimal domains can work for IP-internal A-movement only if it is ensured that a Case position of an argument is never separated from the base position of that argument by an intervening argument position, an assumption that has far-reaching consequences), and Collins (1997, 23).

scrambling of NP_1 to VP in the first step. Due to Fewest Steps (which minimizes optional features on items; cf. section 6 of chapter 2), this movement cannot yet result in checking of [scr], but it is certainly permitted by the (a)-requirement of Last Resort (VP_2-Adj is a typical checking position for the [scr] feature of NP_1). By assumption, VP_2 itself also has a [scr] feature to check. Thus, we end up with the situation that Last Resort can block the first scrambling operation applying to NP_1 in (146) only if γ and β can be identical in the definition of Last Resort in (132), such that VP_2 is closer to *itself* than NP_1. Although I do not really foresee undesirable consequences of such an assumption (but see the remarks in subsection 7.3.2 above), it strikes me as conceptually weird to postulate that closeness can be reflexive (in particular, note that VP cannot actually check off its own [scr] feature). But if we assume that VP_2 does not count as closer to VP_2 than NP_1, Last Resort does not block the derivation in (146), and the illformedness of (145) returns as a problem.

To sum up, we have seen that the MLC as part of Last Resort makes some problematic predictions: In some cases, it may wrongly exclude sentences as completely ungrammatical that either have an intermediate status (cf. the English data in subsection 7.4.1) or are perfectly well formed (cf. the German data in subsection 7.4.2). In other cases, it wrongly permits sentences that could be ruled out as ungrammatical under Unambiguous Domination (cf. (145)). Nevertheless, all in all I believe that a Last Resort approach to the restrictions of remnant movement along the lines of Kitahara (1994), Fox (1995), and Koizumi (1995) is worth pursuing as an alternative to the Unambiguous Domination condition, and that the most problematic consequences of the revised Last Resort condition can be tackled by developing an appropriate concept of closeness – a task that has independently been recognized as important by the consideration of the options of IP-internal Case-driven movement operations.[70]

7.4.4. Bi-Directional Adjunction

Finally, let me point out a potentially positive consequence of the MLC-based notion of Last Resort in (110). Recall from chapter 4 that a specific constraint blocking bi-directional adjunction appeared necessary to exclude certain instances of remnant movement. The constraint is repeated here in (147) (= (94) of chapter 4).

(147) If α dominates β, α-adjunction to γ and β-adjunction to γ cannot be bi-directional.

A derivation that is excluded by this constraint is the one in (148) (= (93) from chapter 4):

[70] Note also that the problems discussed in this section also show up in more or less the same form in Chomsky (1995), irrespective of the issue of remnant movement.

(148) a. *D-Structure Representation*:
– hat der Fritz nicht [$_{\text{VP}}$ [$_{\text{PP}}$ da$_1$ mit]$_2$ gerechnet]$_3$
b. *Scrambling of NP$_1$ to VP*:
– hat der Fritz nicht [$_{\text{VP}}$ da$_1$ [$_{\text{VP}}$ [$_{\text{PP}}$ t$_1$ mit]$_2$ gerechnet]$_3$]$_3$
c. *Extraposition of PP$_2$ to VP*:
– hat der Fritz nicht [$_{\text{VP}}$ [$_{\text{VP}}$ da$_1$ [$_{\text{VP}}$ t$_2$ gerechnet]$_3$]$_3$ [$_{\text{PP}}$ t$_1$ mit]$_2$]$_3$
d. *Scrambling of NP$_1$ to IP*:
– hat da$_1$ der Fritz nicht [$_{\text{VP}}$ [$_{\text{VP}}$ t$'_1$ [$_{\text{VP}}$ t$_2$ gerechnet]$_3$]$_3$ [$_{\text{PP}}$ t$_1$ mit]$_2$]$_3$
e. *Topicalization of VP$_3$*:
[$_{\text{VP}}$ [$_{\text{VP}}$ t$'_1$ [$_{\text{VP}}$ t$_2$ gerechnet]$_3$]$_3$ [$_{\text{PP}}$ t$_1$ mit]$_2$]$_3$ hat da$_1$ der Fritz nicht t$_3$

The intuition behind (147) was that PP$_2$ and NP$_1$ compete for movement to VP in (148), and that PP$_2$-movement to VP blocks NP$_1$-movement since NP$_1$ is dominated by PP$_2$ in the pre-movement structure. This reasoning resembles the MLC-based account of Unambiguous Domination effects given in the preceding subsections – clearly, PP$_2$ is closer to the target than NP$_1$ which is dominated by PP$_2$. Indeed, it seems that (147) can be derived if we assume that extraposition is always triggered by a strong [extr] feature (see subsection 4.1 of chapter 4 and particularly chapter 6 below), and if we adopt a revised version of the Last Resort condition that incorporates a distinction between identical features as interveners and similar features as interveners, as in (149):

(149) *Last Resort*:
α is raised to a position β only if (a) and (b) hold:
a. β is a typical checking position for the lowest-ranked unchecked morphological feature F of α.
b. There is no γ with an unchecked X feature that is closer to β, where X is either identical to F, or X is similar to F and checked in β.

Suppose that one feature is "similar" to another one in the sense relevant for (149-b) if the features are not identical but can be checked in structurally identical positions of the same minimal residue (checking domain). Thus, e.g., [scr] and [extr] are similar. This revision leaves all previous reasonings intact. In addition, (147) follows from Last Resort: In an abstract pre-movement structure like (150)

(150) ... [$_{\text{VP}}$ [$_{\text{PP}_{2[extr]}}$ NP$_{1[scr]}$ P] V] ...

Last Resort now blocks a derivation *NP scrambling to VP* $\succ$ *PP extraposition to VP* if the unchecked [extr] feature of PP$_2$ is checked in the VP-right-adjoined position (but not if PP's [extr] feature is eventually checked in an IP- or CP-adjoined position). In contrast, Last Resort in (149) does not block a derivation *PP extraposition to VP* $\succ$ *NP scrambling to VP*, even if PP checks its [extr] feature in the VP-right-adjoined position: Only a PP with an unchecked [extr] feature can block [scr] feature checking in the same domain by NP. However, this derivation violates the Barriers Condition. Along these lines, it seems that the effects of condition (147) can be made to follow from a Last Resort condition that incorporates the MLC. The price to pay here is that Last Resort is now not formulated in a strictly local way anymore, but has a look-ahead property: To find out whether a movement operation (like NP scrambling to VP in (148)) respects Last Resort, one has to know whether subsequent movement operations (like PP scrambling to VP in (148)) result in actual feature checking or not. Furthermore, it might arguably be preferable on conceptual grounds to dispense with the distinction between identical features and similar features as interveners for the purposes of the MLC; but this distinction is clearly necessary for the revised Last Resort condition to work (thus, it must be ensured that, e.g., extraposition of a PP to IP and scrambling of an NP from the same PP to VP can co-occur, even though PP is closer to both VP and IP in the pre-movement structure). For these reasons, I will only tentatively adopt the revision of Last Resort in this subsection, and draw a general conclusion.

8. Conclusion

In this chapter, I set out to derive an apparent asymmetry between movement types – remnant XPs may undergo topicalization and *wh*-movement, but, in general, they cannot be scrambled or left dislocated. I then went on to call the validity of this statement into question, and proposed instead that the correct empirical generalization is this:

(151) *Generalization*:
Remnant XPs cannot undergo Y-movement if the antecedent of the unbound trace has also undergone Y-movement.

According to this view, any movement type may or may not affect remnant XPs, depending on the position of the antecedent of the unbound trace created by remnant movement: Remnant scrambling is impossible if the antecedent of the unbound trace has also undergone scrambling, but not otherwise; remnant *wh*-movement is impossible if the antecedent of the unbound trace has also undergone *wh*-movement, but not otherwise; etc. In order to account for these facts, I proposed a principle of Unambiguous Domination that precludes (in its representational formulation) α-dominance of α-traces. After refining the analysis in terms of Unambiguous Domination (by considering its relevance for various trace types, and by defining a local domain in which the requirement holds), I then took up recent suggestions that the constraint on remnant movement could be derived from an independently motivated shortest path requirement, the MLC; and I located this requirement within the Last Resort condition. An investigation into the predictions of such a revised approach has shown it to be capable of deriving the basic effects of Unambiguous Domination; but whether or not it can be maintained in the light of certain pieces of counter-evidence is, to my mind, an open question that awaits further research.

Chapter 6

Intermediate Traces and Extraposition

1. The Problem

In this chapter, I would like to address the last of the four problems raised by remnant movement that were noted in section 3 of chapter 1, namely, the apparent inability of remnant movement to create unbound intermediate traces (cf. den Besten & Webelhuth (1990), Fanselow (1992; 1993), Grewendorf (1994), and Bayer (1996)). Two relevant constructions that initially lend support to this view are repeated here in (1) and (2) (= (63) and (65-a) from chapter 1):

(1) *[$_{\text{CP}}$ t_1' Daß Fritz t_1 liebt]$_2$ weiß ich nicht [$_{\text{CP}}$ wen$_1$ er t_2 gesagt
that Fritz$_{nom}$ loves know I$_{nom}$ not who$_{acc}$ he$_{nom}$ said
hat]
has

(2) *[$_{\text{VP}}$ t_2 Gesagt [$_{\text{CP}_2}$ t_1' dass Fritz t_1 liebt]]$_3$ weiß ich nicht [$_{\text{CP}}$ wen$_1$
said that Fritz$_{nom}$ loves know I$_{nom}$ not who$_{acc}$
er t_3 hat]
he$_{nom}$ has

In (1) successive-cyclic *wh*-movement (creating an intermediate trace t_1' for reasons of locality) occurs together with topicalization of the remnant CP, and the result is ill formed. Similarly, in (2) successive-cyclic *wh*-movement and remnant topicalization co-occur; but in this case, what is topicalized is a remnant VP that includes an extraposed CP that contains the traces of *wh*-movement. These constructions both involve a *wh*-island effect, but they contrast sharply with their counterparts in (3) and (4), where no remnant movement occurs.

(3) ??[$_{\text{CP}}$ Daß Fritz Caroline liebt]$_2$ weiß ich nicht [$_{\text{CP}}$ ob er t_2
that Fritz$_{nom}$ Caroline$_{acc}$ loves know I$_{nom}$ not whether he$_{nom}$
zugeben würde]
admit would

(4) ??[$_{\text{VP}}$ t$_2$ Gesagt [$_{\text{CP}_2}$ dass Fritz Caroline liebt]]$_3$ weiß ich nicht [$_{\text{CP}}$
said that Fritz$_{nom}$ Caroline$_{acc}$ loves know I$_{nom}$ not
ob er t$_3$ hat]
whether he$_{nom}$ has

These cases of topicalization of a non-remnant XP across a *wh*-island have an intermediate status; i.e., they are much better than the remnant movement cases in (1) and (2). Thus, there seems to be an additional factor involved that renders (1) and (2) sharply ungrammatical.[1]

It is tempting to assume that the underlying generalization is that intermediate unbound traces may never occur and that (5) holds as a principle:[2]

(5) Unbound traces must be initial.

However, there are a number of factors that call the validity of (5) into question. I address these in the following section.

2. Remnant Movement Can Generate Intermediate Traces

2.1. A Conceptual Problem

First, there is a conceptual problem with (5). This condition explicitly differentiates between bound and unbound traces. Hence, (5) can be viewed as a construction-specific assumption of exactly the type I have sought to avoid in the preceding chapters – the basic task so far has been to account for the properties of remnant movement

---

[1] At first sight, one might assume that ungrammatical sentences like (i-a) and (i-b) in German pose the same kind of problem as (1) and (2). Here the antecedent of the unbound intermediate trace is in a scrambling position (and not in a *wh*-position). (For (i-b) see Hinrichs & Nakazawa (1994, 44), who observe that this construction is potentially problematic for an approach to incomplete category fronting in terms of remnant movement or, in their analysis, SLASH features on the fronted VP.)

(i) a. *[$_{\text{CP}}$ t$'_1$ Daß Fritz t$_1$ liebt]$_2$ habe ich sie$_1$ t$_2$ gesagt
that Fritz$_{nom}$ loves have I$_{nom}$ her$_{acc}$ said

b. *[$_{\text{VP}}$ t$_2$ Gesagt [$_{\text{CP}_2}$ t$'_1$ dass Fritz t$_1$ liebt]]$_3$ habe ich sie$_1$
said that Fritz$_{nom}$ loves have I$_{nom}$ her$_{acc}$

However, according to the approach adopted here, (i-a) and (i-b) are directly ruled out, irrespective of the issue of remnant movement. The reason is that a derivation of these sentences must involve long-distance scrambling from a finite clause, which is blocked by the interaction of the Barriers Condition and Last Resort (cf. chapter 2).

[2] (5) builds on an insight by den Besten & Webelhuth (1990, 83 & 86) but is not identical to either of the two generalizations envisaged by these authors. They first assume that the apparent restriction blocks unbound traces in specifier positions, and eventually suggest that all traces in non-argument positions are precluded from occurring unbound. I will later discuss some of the evidence that leads den Besten & Webelhuth to these assumptions; for the time being, recall from chapter 1 that it looks as though non-argument traces can indeed show up unbound with remnant topicalization in German.

constructions without actually stipulating conditions that refer to the bound/unbound status of a trace. Thus, to the extent that (5) correctly describes the situation in (1) and (2), this should ideally follow from some general, non-construction specific principle(s).

2.2. A Redundancy

Second, it was shown in chapter 4 (subsection 5.2) that one of the two kinds of examples just mentioned that seem to support (5) can be accounted for independently, as involving a Freezing effect. This is (2), which is repeated here:

(6) *[$_{\text{VP}}$ t$_2$ Gesagt [$_{\text{CP}_2}$ t$'_1$ dass Fritz t$_1$ liebt]]$_3$ weiß ich nicht [$_{\text{CP}}$ wen$_1$
said that Fritz$_{nom}$ loves know I$_{nom}$ not who$_{acc}$
er t$_3$ hat]
he$_{nom}$ has

This sentence necessarily involves three movement operations: *wh*-movement, VP topicalization, and CP extraposition. *Wh*-movement must apply first; otherwise, the Barriers Condition would be violated. If CP extraposition takes place second, before VP topicalization, the Strict Cycle Condition is violated. And if VP topicalization applies second, CP extraposition in the third step violates the Strict Cycle Condition. Thus, the illformedness of examples like (6) follows directly from the derivational approach to (Anti-) Freezing developed in chapter 4, and resorting to (5) is not necessary. If one nevertheless wants to maintain (5) in light of this, it is clear that massive redundancies arise that one might want to avoid.

2.3. A Missed Generalization

What is more, it was demonstrated in subsections 5.2.2 and 5.2.3 of chapter 4 that exactly the same Freezing effect as in (6) shows up with extraposition of a PP and of a coherent infinitive (the third construction), as in the following examples:

(7) *[$_{\text{VP}}$ [$_{\text{VP}}$ t$_2$ Gerechnet]$_3$ [$_{\text{PP}}$ t$_1$ mit]$_2$]$_3$ hat da$_1$ der Fritz nicht t$_3$
counted with has there ART Fritz not

(8) *[$_{\text{VP}}$ t$_2$ Versucht [$_\alpha$ dem Peter t$_1$ zu geben]$_2$]$_3$ hat die Claudia das
tried ART Peter$_{dat}$ to give has ART Claudia the
Buch$_1$ nicht t$_3$
book$_{acc}$ not

The ungrammaticality of (7) and (8) follows along the same lines as that of (6). However, these examples do not involve intermediate traces in the moved remnant XPs. As noted in chapter 1, extraction of an R-pronoun from a PP and from a coherent infinitive may take place without establishing an intermediate (PP-internal or infinitive-internal) trace; and in line with this, PP-internal traces and traces included in coherent infinitives may in principle occur unbound, as a result of remnant movement. Thus, it seems intuitively appealing to find a common reason for the ungrammaticality of (6) on the one hand and of (7) and (8) on the other. This is accomplished under the approach to Freezing in chapter 4 but not under the condition (5), which cannot account for the prohibition against (7) and (8).

2.4. Unbound Intermediate Traces in Extraposed CPs

Fourth, as was noted in subsection 5.2.1 of chapter 4, there is an empirical problem with the condition in (5) that arises under the approach to extraposition developed there. Recall that the explanation of the Anti-Freezing effect with CP extraposition in German that was given in chapter 4 crucially relies on the idea that *wh*-movement can precede CP extraposition in a sentence like (9) (as demanded by the Barriers Condition) without violating the Strict Cycle Condition only if CP extraposition ends up in a CP-adjoined position, i.e., if CP extraposition in (9) is an instance of remnant movement, creating an unbound intermediate *wh*-trace t'_1, in violation of (5):

(9) (Ich weiß nicht) wen$_1$ er t_2 gesagt hat [$_{CP}$ t'_1 dass Claudia t_1 geküßt hat]$_2$
I know not whom he said has that Claudia kissed has

One might object that the evidence to be gained against (5) from (9) depends on a theory-internal assumption, viz., that CP here is indeed in a higher position at S-structure than the *wh*-phrase *wen*$_1$ (‘whom’), so that t'_1 is unbound. No such objection would be possible if well-formed examples could be found that are structurally analogous to the ones in (1) and (2) in the sense that an intermediate unbound trace is unquestionably present. I turn to this in the following subsection.

2.5. Unbound Intermediate Traces in Topicalized Infinitives

Incoherent infinitives behave just like finite clauses in German with respect to an extraction asymmetry between scrambling and specifier movement operations like *wh*-movement. Thus, a finite object CP in German may permit *wh*-extraction (if it is embedded by a bridge verb), but it never permits long-distance scrambing out of it (cf. subsection 4.2 of chapter 2 and the references given there):

(10) a. *dass Antje Hygrometer$_1$ sagt [$_{CP}$ t'_1 dass niemand t_1 mag]
that Antje hygrometers says that no-one likes

b. Was$_1$ denkt Antje [$_{CP}$ t'_1 dass niemand t_1 mag] ?
What thinks Antje that no-one likes

As was briefly mentioned in footnote 17 of chapter 1, the same asymmetry shows up with incoherent object infinitives like those embedded by *ablehnen* (‘refuse’) in German. Consider the following contrast:

(11) a. ?*dass [$_{NP}$ das Buch]$_1$ keiner [$_\alpha$ t$_1$ zu lesen] abgelehnt hat
that the book$_{acc}$ no-one$_{nom}$ to read refused has

b. Was$_1$ hat keiner [$_\alpha$ t$_1$ zu lesen]$_3$ abgelehnt ?
what$_{acc}$ has no-one$_{nom}$ to read refused

It is commonly assumed that incoherent infinitives (in contrast to coherent infinitives) retain the full clausal structure and the transparency properties of finite clauses, i.e., that α in (11) is a CP. If this is so, the asymmetry between scrambling and *wh*-movement in (11) follows in the same way as that in (10). As regards (10), I argued in chapter 2 (following Müller & Sternefeld (1993)) that the embedded CP is a barrier that can only be circumvented by an intermediate movement step to SpecC. Such an operation is possible if the ultimate landing site is SpecC, too; however, if long-distance scrambling uses this option, a PUB effect arises that has been shown to be derivable from economy (long-distance scrambling of *Hygrometer* via the embedded SpecC position violates Last Resort because the scrambled item is not equipped with a [+wh] feature). Evidently, under the assumption that incoherent infinitives are complete CPs, extraction from such an infinitive depends on the availability of SpecC as an escape hatch, and we expect a PUB effect induced by Last Resort in the case of long-distance scrambling in (11-a) but not in the case of *wh*-movement in (11-b). What is relevant in the present context, however, is not the prohibition against scrambling from an incoherent infinitive as such, but rather the fact that this prohibition shows that well-formed cases of extraction from such an infinitive, as in (11-b), must proceed via SpecC, i.e., must create an intermediate trace:

(12) Was$_1$ hat keiner [$_{CP}$ t$'_1$ C PRO t$_1$ zu lesen]$_3$ abgelehnt ?
what$_{acc}$ has no-one$_{nom}$ to read refused

Exactly the same conclusion can be drawn in the case of subject infinitives, which are necessarily incoherent in German. Again, these infinitives do not permit long-distance scrambling, but they permit *wh*-extraction (cf. the discussion in Haider (1983), Sternefeld (1985a), and Grewendorf (1989)); and again, this can be taken to indicate that they come equipped with a full CP structure and require extraction to proceed via SpecC.[3] Consider the examples in (13) (here, (13-a) involves a subject infinitive without extraction, preceded by a sentence adverbial of the matrix clause so as to mark the clause boundary of the infinitive; (13-b) shows that the object of the subject infinitive cannot be scrambled into the matrix clause, in front of the sentence adverbial; and (13-c) makes the point that *wh*-extraction is possible in this context):

[3] As noted before (see footnote 5 of chapter 2), the very possibility of *wh*-extraction from an infinitival subject CP in German is unexpected under the approach to barriers adopted here. To my mind, this fact ultimately suggests that a different type of barriers theory (in terms of minimality) should be developed. But, as before, I will not attempt to do such a thing here; let us simply assume that the general transparency of infinitivals for *wh*-extraction (topicalization, relativization) in German can be derived from the theory of locality in one way or another.

(13) a. dass wahrscheinlich [$_{\text{CP}}$ [$_{\text{NP}}$ diesen Film]$_1$ zu beanstanden]
that probably this film to object to
sich nicht gehören würde
not be proper would

b. *dass [$_{\text{NP}}$ diesen Film]$_1$ wahrscheinlich [$_{\text{CP}}$ (t'_1) PRO t_1 zu beanstanden]
that this film probably to object to
sich nicht gehören würde
not be proper would

c. Was$_1$ würde wahrscheinlich [$_{\text{CP}}$ t'_1 PRO t_1 zu beanstanden]
what would probably to object to
sich nicht gehören ?
not be proper

Since incoherent object and subject infinitives do not allow long-distance scrambling in the first place due to the interaction of the Barriers Condition and Last Resort, we of course expect them not to be able to contain unbound scrambling traces after remnant movement (e.g., remnant topicalization) either. This was already shown for incoherent object infinitives in subsection 2.5 of chapter 1 (cf. (42-b) from that chapter, repeated here as (14-a)) and is shown for subject infinitives in (14-b):

(14) a. *[$_{\text{CP}}$ (t'_1) t_1 Zu lesen]$_3$ hat [$_{\text{NP}}$ das Buch]$_1$ keiner t_3 abgelehnt
to read has the book no-one refused

b. *[$_{\text{CP}}$ (t'_1) t_1 Zu beanstanden]$_3$ würde [$_{\text{NP}}$ diesen Film]$_1$ sich nicht gehören
to object to would this film not be proper

However, an interesting question that now arises is how incoherent infinitives behave with respect to unbound traces of *wh*-movement. By analogy with finite clauses, we might expect the pertinent constructions to be utterly impossible, and this is of course what the condition in (5) predicts, given that *wh*-extraction from these kinds of clauses must establish an intermediate trace in SpecC in order to be possible at all. The data in (15), though, show that this prediction is wrong – unbound intermediate traces of *wh*-movement may show up in topicalized incoherent infinitives ((15-a) involves topicalization of an incoherent object infinitive, and (15-b) shows the same with topicalization of an incoherent subject infinitive):[4]

[4] A terminological remark is in order here. Up to now, usage of the notion "remnant infinitive" has implied that the infinitive is coherent, i.e., permits scrambling out of it. However, now it turns out that remnant infinitives may also be incoherent infinitives, viz., if unbound *wh*-traces are involved. Hence, remnantness of an infinitive does not automatically signal coherence anymore.

(15) a. ??[$_{CP}$ t$'_1$ PRO t$_1$ Zu lesen]$_3$ weiß ich nicht [$_{CP}$ was$_1$ der Fritz t$_3$ abgelehnt
to read know I not what ART Fritz refused
hat]
has

b. ??[$_{CP}$ t$'_1$ PRO t$_1$ Zu beanstanden]$_3$ weiß ich nicht [$_{CP}$ was$_1$ t$_3$
to object to know I not what
sich nicht gehört]
is not proper

These examples involve the weak *wh*-island effect that is typical of topicalization of arguments across a *wh*-item in SpecC, but they do not exhibit the severe deviance that can be observed with analogous cases involving finite clauses, as in (1). Thus, (15-a) and (15-b) are completely on a par with the corresponding sentences in (16) that do not involve unbound traces in the fronted infinitives:

(16) a. ??[$_{CP}$ Das Buch zu lesen]$_3$ weiß ich nicht [$_{CP}$ warum der Fritz t$_3$
the book to read know I not why ART Fritz
abgelehnt hat]
refused has

b. ??[$_{CP}$ Diesen Film zu beanstanden]$_3$ weiß ich nicht [$_{CP}$ ob t$_3$
this film to object to know I not whether
sich nicht gehört]
is not proper

To sum up so far: However the contrast between intermediate unbound traces in topicalized finite clauses (as in (1)) and intermediate unbound traces in topicalized non-finite clauses (as in (15-a) and (15-b)) ultimately is to be accounted for, the mere existence of the sentences in (15) shows that the condition in (5) cannot be correct – unbound intermediate traces may in principle occur in topicalized CPs.

What, then, is responsible for the impossibility of remnant movement in cases like (1), if not a general prohibition againt unbound intermediate traces such as (5)? Towards a solution, I will now first turn to infinitival counterparts of the sentence in (2).

2.6. Unbound Intermediate Traces With and Without Extraposition

Consider first an example like (17):

(17) ?*[$_{VP_2}$ [$_{VP_2}$ t$_3$ Abgelehnt] [$_{CP_3}$ t$'_1$ dem Peter t$_1$ zu geben]] weiß ich nicht
refused ART Peter$_{dat}$ to give know I not
[$_{CP}$ was$_1$ er t$_2$ hat]
what$_{acc}$ he has

This sentence is structurally analogous to (2), (7), and (8), and like these constructions, it is ungrammatical. This conforms to expectations. Due to the Barriers Condition, *wh*-movement of *was*$_1$ ('what') to its target position must apply first, and no matter in which order the other two movement operations (VP topicalization and CP extraposition) apply, the Strict Cycle Condition is invariably violated.[5] With this in mind, consider (18):

(18) ??[$_{VP_2}$ [$_{CP_3}$ t'_1 Dem Peter t_1 zu geben] abgelehnt] weiß ich nicht [$_{CP}$
ART Peter$_{dat}$ zu geben refused know I not
was$_1$ er t_2 hat]
what$_{acc}$ he has

This sentence differs minimally from that in (17) in that VP-internal extraposition of CP has not taken place, and the result is not worse than any *wh*-island violation with topicalization in German; i.e., it is better than (17). Again, this does not come as a surprise under present assumptions – a derivation of (18) in which *wh*-movement precedes VP topicalization does not violate any constraints (in particular, it does not violate the Strict Cycle Condition). Clearly, the prediction now is that – ceteris paribus – the same conclusions should apply if CP, in a construction that takes the form of (18), were finite and not an infinitival clause. However, it is *not* the case that all other things are equal. There is an intervening factor with finite clauses, and this factor was already noted in subsection 4.1 of chapter 4 and left as an open problem there: Finite clauses, in contrast to infinitives (of all kind, i.e., both coherent and incoherent nonfinite clauses), permit *wh*-extraction (or any other kind of long-distance movement) only if they eventually show up in an extraposed position. Thus, sentences like (19-a) (= (58-a) from chapter 4), in which the CP from which *wh*-extraction has taken place remains in situ, are ungrammatical; and so are sentences like (19-b) (which is structurally analogous to (18) and differs minimally from (2) in that CP extraposition has not taken place), for the very same reason:

(19) a. *(Ich weiß nicht) wen$_1$ er [$_{CP}$ t'_1 dass Claudia t_1 geküßt hat]$_2$ gesagt hat
I know not whom he that Claudia kissed has said has

b. *[$_{VP}$ [$_{CP_2}$ t'_1 Daß Fritz t_1 liebt] gesagt]$_3$ weiß ich nicht [$_{CP}$ wen$_1$
that Fritz$_{nom}$ loves said know I$_{nom}$ not who$_{acc}$
er t_3 hat]
he$_{nom}$ has

Minimally contrasting examples involving *wh*-extraction from infinitives are given in (20) (also see (12)) :

(20) a. (Ich weiß nicht) wen$_1$ er [$_{CP_2}$ t'_1 C PRO t_1 zu küssen] abgelehnt hat
I know not whom$_{acc}$ he to kiss refused has

[5] It might be the case, however, that the degree of ungrammaticality in the case of (17) is not quite the same as with analogous constructions like those just mentioned; this is indicated in the example by assigning it a "?*" instead of a "*". At present, I have no explanation for this subtle difference.

b. (Ich weiß nicht) wen$_1$ er t$_2$ abgelehnt hat [$_{CP_2}$ t$'_1$ C PRO t$_1$ zu
I know not whom$_{acc}$ he refused has to
küssen]
kiss

Also recall that examples with a finite CP in situ are much better (albeit often far from perfect, for reasons which are unclear to me) if *wh*-extraction does not occur; compare (19-a) and (19-b) with (21-a) (= (57-a) from chapter 4) and (21-b):

(21) a. ??dass er [$_{CP}$ dass Claudia Peter geküßt hat]$_2$ gesagt hat
that he that Claudia Peter kissed has said has

b. ??[$_{VP}$ [$_{CP}$ Daß Claudia Peter geküßt hat]$_2$ gesagt]$_3$ hat er t$_3$ nicht
that Claudia Peter kissed has said has he not

I would like to contend that the conclusion to be drawn from all this is the following: The ungrammaticality of sentences like (1) has nothing to do with the fact that an intermediate unbound trace shows up; rather, it should be related to the observation that extraction from a finite CP in German seems to necessarily require extraposition of that CP, as indicated by the illformedness of sentences like those in (19). According to this view, what saves sentences involving topicalization of remnant infinitives (or categories containing remnant infinitives) that include unbound intermediate traces (as in (15) and (18)) is the fact that extraction from a non-finite CP does not have to go hand in hand with extraposition of that CP. In the following section, I will develop this idea.

3. Extraction as a Trigger for Extraposition

3.1. Extraposition and Last Resort

Let me begin by considering the consequences that the Last Resort condition has for extraposition more closely. The first thing to note is that given Last Resort, all kinds of extraposition (clausal extraposition, Heavy NP Shift, PP extraposition, etc.) must be triggered by an appropriate feature. Indeed, I have (more or less tacitly) assumed so far that in all the cases of extraposition that were considered, an [extr] feature is indeed present, in the pre-movement structure, on the item that is to undergo extraposition; recall footnote 18 of chapter 4 and subsection 7.4.4 of chapter 5. This [extr] feature behaves in many respects like the [top] and [scr] features assumed for topicalization and scrambling, respectively. As in the case of these latter features, it seems plausible to assume that the [extr] feature is [–Interpretable] and necessarily strong if it is present at all (i.e., a concept of LF extraposition appears to be poorly motivated); moreover, I would like to postulate that the [extr] feature is in principle optional, just like [scr] and [top]. As for the target of extraposition, I will simply continue to assume that any right-adjoined position in the minimal residue of a head can function as a target for extraposition, i.e., that any head can check off the [extr] feature on a lexical item; cf.

(66) from chapter 4.

Abstracting away from a few potentially intervening factors, I take the assumption that [extr] is optionally inserted to be fairly uncontroversial with respect to PPs in SOV languages like German.[6] As for NPs, assuming [extr] to be a strictly optional feature might appear problematic at first sight. After all, only heavy NPs can undergo extraposition in German, and given that the [extr] feature shows up on bare nouns in the numeration, it seems that we are deprived of a means of relating the presence or absence of an [extr] feature on N to the "heaviness" of the projection headed by N. However, this problem disappears if we follow Rochemont (1978), Fanselow (1991), and others in assuming that the decisive factor in licensing NP extraposition is not actually heaviness, but rather focussing. This latter property can be marked inherently on a lexical head in the numeration, and we can stipulate that the [extr] feature, while in principle optional, can be instantiated on a noun only if that noun is also marked as focussed.[7]

This leaves us with CPs. As for control infinitives, assuming [extr] to optionally show up on an infinitival C makes the correct predictions; cf.:[8]

(22) a. dass Fritz [$_{\text{CP}_1}$ C_1 das Buch zu lesen] versucht/abgelehnt hat
that Fritz the book to read tried/refused has

b. dass Fritz t_1 versucht/abgelehnt hat [$_{\text{CP}_1}$ C_1 das Buch zu lesen]
that Fritz tried/refused has the book to read

As far as finite CPs in German are concerned, I have assumed that exactly the same situation holds, i.e., that extraposition is by itself not obligatory (leaving it open whether – and if so, how – the slight deviance or "clumsiness" of middle field-internal finite CPs in German should be accounted for in syntax):

---

[6]It should be noted, though, that PP extraposition – perhaps somewhat surprisingly – yields better results with light PPs than with heavy PPs; cf.:

(i) a. dass er ihn t_1 gebeten hat darum$_1$
that he$_{nom}$ him$_{acc}$ asked has for that

b. ?dass er ihn t_1 gebeten hat [$_{\text{PP}_1}$ um diesen Gefallen]
that he$_{nom}$ him$_{acc}$ asked has for this favour

Other factors that are potentially relevant in this domain involve intonation and the grammatical function of the extraposed PP. In what follows, I abstract away from these factors, assuming that they do not reflect genuine syntactic restrictions of the type that should be captured by constrains on [extr] insertion.

[7]I.e., what we have here is a GPSG-type feature co-occurrence restriction, as discussed earlier in chapter 2. As before, one would ideally like to derive this from more general assumptions, but I will not attempt to do that here.

[8]The case is different with raising and ECM infinitives, where extraposition is (generally) impossible in German. I will not address this problem here.

(23) a. ??dass er [$_{CP}$ dass$_1$ Claudia Peter geküßt hat]$_1$ gesagt hat
that he that Claudia Peter kissed has said has

b. dass er t$_1$ gesagt hat [$_{CP_1}$ dass$_1$ Claudia Peter geküßt hat]
that he said has that Claudia Peter kissed has

In (22-a) and (23-a), an [extr] feature has failed to show up on C$_1$ in the numeration; hence, extraposition does not occur. In (22-b) and (23-b), in contrast, there is an optional [extr] feature on the C$_1$ node, and extraposition takes place, in line with Last Resort.

Needless to say, these considerations do not yet comprehensively account for all the possibilities and restrictions related to extraposition in a language like German, but they may suffice for present purposes.

3.2. Extraction and Extraposition

3.2.1. Derivational [extr] Instantiation

Now recall that *wh*-extraction from a finite CP is possible in German only if that CP ends up in an extraposed position. Cf. the contrast between (9) and (19-a); both examples are repeated here:

(24) a. (Ich weiß nicht) wen$_1$ er t$_2$ gesagt hat [$_{CP}$ t$'_1$ dass Claudia t$_1$ geküßt hat]$_2$
I know not whom he said has that Claudia kissed has

b. *(Ich weiß nicht) wen$_1$ er [$_{CP}$ t$'_1$ dass Claudia t$_1$ geküßt hat]$_2$ gesagt hat
I know not whom he that Claudia kissed has said has

This contrast does not show up only in German; a cross-linguistic survey reveals that it is quite a common property of finite complements in SOV languages that they permit overt *wh*-extraction only under extraposition. For instance, as has been noted by Ormazabal, Uriagereka, & Uribe-Etxebarria (1994, 6-7), the situation in Basque is completely parallel. Thus, it seems that what we want to express is that an otherwise optional feature (viz., [extr]) is instantiated obligatorily on a finite C node if extraction takes place via the minimal residue of C. I take it that this cannot be expressed in strictly local terms in an insightful way – in the numeration there is nothing that tells us whether the minimal residue of a [–wh] C node is later going to be affected by *wh*-extraction across it or not. Thus, I would like to suggest adopting an idea that was first (to my knowledge) proposed by Watanabe (1993, 56ff). Watanabe assumes that features that must be checked off by movement are not necessarily present in the numeration. Rather, some features may arise as a consequence of syntactic movement (and checking) operations. This holds, e.g., for Case-driven movement; according to Watanabe's assumptions, NP-movement to SpecAGR creates a new feature [F] that must be checked off by further movement to a higher head. Elaborating on this general idea, I will assume that the [extr] feature may indeed be instantiated on a C node in one of two ways – either by optional insertion in the process of creating the numeration or by *wh*-extraction in the derivation.[9] This latter way of instantiating an [extr] feature is explicitly stated in (25):[10]

(25) *Derivational [extr] Instantiation*:
An $[_{\text{extr}]}$ feature is instantiated on a finite non-root C if the minimal residue of this C node is affected by movement.

From Derivational [extr] Instantiation it follows that (24-b) is ungrammatical. Here, *wh*-extraction from CP instantiates an [extr] feature on C, but CP fails to undergo extraposition, so that the [extr] feature remains unchecked. Hence, the derivation that generates this S-structure representation does not converge. In (24-a), on the other hand, the [extr] feature instantiated on C is checked off by CP extraposition; the derivation converges, and the sentence is well formed.

A technical question arises in this context. Recall that the [extr] feature is in principle optional on C. So, what happens if C comes equipped with one [extr] feature assigned to it in the numeration, and another one acquired in the derivation? (25) would rule out such a situation if it were accompanied by a statement that C can receive an [extr] feature derivationally only if it is [extr]-less so far. However, it turns out that such a stipulation is unneccessary under the revised notion of Fewest Steps that was proposed in section 6 of chapter 2 (in (113), repeated here as (26)):

(26) *Fewest Steps*:
If two derivations D_1 and D_2 are in the same reference set and D_1 involves fewer checking operations than D_2, then D_1 is to be preferred over D_2.

If a C node has two [extr] features, these two features may well be checked off on one and the same target head. However, in that case, there are two checking operations in which the derivation participates. A competing derivation in which C does not initially bear an [extr] feature in the derivation will therefore block the derivation just envisaged, in which C ends up having two [extr] features at a certain point in the derivation. Thus, again (i.e., as in the cases of illegitimate optional movement discussed in section 6 of chapter 2), the revised Fewest Steps condition has the effect of minimizing

[9] Like Watanabe's (1993) proposal, this view is not compatible with Chomsky's (1995, 228) inclusiveness condition according to which "no new objects are added in the course of computation."

[10] For the time being, I assume that (25) holds in German, Basque, and other SOV languages; I will abstract away from a possible extension to other language types. – Note that the confinement to non-root C nodes in (25) is necessary because an [extr] feature on a root C node affected by *wh*-movement would yield the disastrous result that all *wh*-movement should be impossible in root clauses; the reason for this is that an [extr] feature on a root C node could never be checked off anymore.

optional features.[11]

As it stands, (25) not only triggers finite CP extraposition if *wh*-movement crosses the minimal residue of C, ending up in a higher position in the clause, but also predicts that CP extraposition must apply if the minimal residue of C acts as the ultimate landing site of a *wh*-movement operation. In other words, finite embedded questions formed by *wh*-movement always require extraposition of CP. I think this consequence is indeed tenable; cf. the contrast in (27):[12]

(27) a. *dass Fritz sich [$_{\mathrm{CP}_2}$ warum$_1$ C sie t$_1$ das getan hat] fragt
that Fritz$_{nom}$ REFL why she$_{nom}$ that$_{acc}$ done has asks

b. dass Fritz sich t$_2$ fragt [$_{\mathrm{CP}_2}$ warum$_1$ C sie t$_1$ das getan hat]
that Fritz$_{nom}$ REFL asks why she that done has

The in situ version in (27-a) seems to be more on a par with a sentence like (24-b) than with a sentence like (23-a), as far as its degree of (un-) grammaticality is concerned, and this follows under the formulation of (25) just given. As one might expect, the same strong contrast arises with VP topicalization:

(28) a. *[$_{\mathrm{VP}_3}$ [$_{\mathrm{CP}_2}$ Warum$_1$ C sie t$_1$ das getan hat] gefragt] hat Fritz sich nicht t$_3$
why she that done has asked has Fritz REFL not

b. [$_{\mathrm{VP}_3}$ t$_2$ Gefragt [$_{\mathrm{CP}_2}$ warum$_1$ C sie t$_1$ das getan hat]] hat Fritz sich nicht t$_3$
asked why she that done has has Fritz REFL not

3.2.2. Topicalization

Closer inspection reveals that the formulation of Derivational [extr] Instantiation given in (25) is not yet general enough. Like *wh*-movement, topicalization from an embedded finite clause in German is possible only if that clause undergoes extraposition; cf.:

(29) a. Den Karl$_1$ hat sie t$_2$ gesagt [$_{\mathrm{CP}_2}$ t$'_1$ dass sie t$_1$ geküßt hat]
ART Karl$_{acc}$ has she said that she kissed has

---

[11] Note that the Fewest Steps condition allows optional extraposition in the first place for the same reason that it allows optional scrambling: The LF representation of a derivation in which an optional [extr] feature is checked is different from an LF representation of an otherwise identical derivation in which no [extr] feature (hence, no extraposition) is present.

[12] However, (25) could straightforwardly be reformulated in such a way that it triggers CP extraposition only if movement proceeds further. This is purely an empirical question.

b. *Den Karl$_1$ hat sie [$_{CP_2}$ t$'_1$ dass sie t$_1$ geküßt hat] gesagt
ART Karl$_{acc}$ has she that she kissed has said

Given that topicalization is driven by a [top] feature on an XP that is checked off by the functional category Top (as suggested in chapter 2 and assumed throughout), it seems necessary to extend (25) in such a way that it also accounts for the contrast in (29). This can be accomplished by modifying (25) as in (30):

(30) *Derivational [extr] Instantiation* (revised):
An [extr] feature is instantiated on a finite non-root clausal functional category X if the minimal residue of X is affected by movement.

"Clausal functional category" here is a cover term for C and Top. Note that as it stands (30) does not actually trigger CP extraposition if Top's minimal residue is affected by movement (as in (29)); rather, what is extraposed on this view is TopP. This view might prove to be correct; alternatively, matching of CP and TopP (as had to be postulated anyway, for reasons of locality; recall subsection 4.2.5 of chapter 2) might also ensure that an [extr] feature of Top can be checked against a target head X even if it is not TopP itself that ends up in the minimal residue of X, but rather the CP that immediately dominates (and matches with) TopP. For lack of decisive empirical evidence, I will leave this question open.

In addition to accounting for (29-b), the revised condition of Derivational [extr] Instantiation in (30) makes an interesting prediction concerning topicalization and V/2 clauses in German. As is well known, embedded clauses in which topicalization (and V/2 movement) has applied may never remain in situ in German; see, e.g., Stechow & Sternefeld (1988, 392 & 397-399). Examples that show this restriction are given in (31) (compare the minimally different examples without internal topicalization in (23)):

(31) a. *dass er [$_{CP}$ den Peter$_2$ habe die Claudia t$_2$ geküßt]$_1$ sagte
that he ART Peter$_{acc}$ has$_s$ub ART Claudia$_{nom}$ kissed said

b. dass er t$_1$ sagte [$_{CP_1}$ den Peter$_2$ habe die Claudia t$_2$ geküßt]
that he said ART Peter$_{acc}$ has$_{subj}$ ART Claudia$_{nom}$ kissed

The illformedness of (31-a) follows directly under present assumptions. Both topicalization and V/2 movement affect the minimal residue of the finite clausal functional head Top (in contrast to what is the case with complementizer-initial clauses, as in (23), where the domain of C is not affected by any movement operation). Hence, an [extr] feature is instantiated on Top but not checked off via clausal extraposition, and therefore the derivation generating (31-a) necessarily fails to converge. No such problem arises in (31-b). As before, Top obligatorily receives an [extr] feature (due to topicalization and V/2 movement), but this time subsequent extraposition applies, thereby checking off this feature and leading to convergence. Thus, the prohibition against embedded V/2 clauses in situ in German is derived.

3.3. Intermediate Traces in Topicalized CPs Revisited

3.3.1. The Analysis

With all this mind, we can now return to the original problem posed at the beginning of this chapter, which is to account for the severe illformedness of sentences like (32) (= (1)):

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

The pre-movement representation of this sentence is given in (33) (as before, I abstract away from V/2 movement in the matrix clause, since it is irrelevant to the issue currently under consideration):

(33) – weiß ich [$_{\text{CP}_3}$ – er [$_{\text{CP}_2}$ – dass Fritz wen$_1$ liebt] gesagt hat] nicht

To fulfill the Barriers Condition, the first two movement operations must be *wh*-extractions applying to *wen*$_1$ and moving it via SpecC$_2$ to SpecC$_3$:

(34) a. – weiß ich [$_{\text{CP}_3}$ – er [$_{\text{CP}_2}$ wen$_1$ dass Fritz t_1 liebt] gesagt hat] nicht
b. – weiß ich [$_{\text{CP}_3}$ wen$_1$ er [$_{\text{CP}_2}$ t'_1 dass Fritz t_1 liebt] gesagt hat] nicht

Substitution of *wen*$_1$ in SpecC$_2$ creates an [extr] feature on C$_2$, given Derivational [extr] Instantiation in (30), and, for the same reason, subsequent substitution of *wen*$_1$ in SpecC$_3$ creates an [extr] feature on C$_3$. Now we know that three additional movement operations must take place in order to generate the S-structure representation in (32), viz., CP$_2$ extraposition, CP$_3$ extraposition, and CP$_2$ topicalization. Because of the Strict Cycle Condition, CP$_2$ extraposition to the minimal residue of C$_3$ must apply next, as in (35):

(35) – weiß ich [$_{\text{CP}_3}$ [$_{\text{CP}_3}$ wen$_1$ er t_2 gesagt hat] [$_{\text{CP}_2}$ t'_1 dass Fritz t_1 liebt]] nicht

The [extr] feature on C$_2$ is checked off by this movement operation. Next, we can assume that CP$_3$ undergoes extraposition, e.g., to IP (it may also undergo extraposition to the matrix CP/TopP). And finally, topicalization applies to CP$_2$ (which implies that C$_2$ has been assigned an optional [top] feature in the numeration), thereby creating the S-structure representation in (32). (Since neither of these two last movement operations is in any way problematic as regards locality conditions, and since both operations may end up in the very same minimal residue so that the Strict Cycle Condition does not necessarily impose a specific order, it might just as well be the case that the order of rule application in (36) is reversed. Nothing depends on this in the present context.)

(36) a. – weiß ich t_3 nicht [$_{\text{CP}_3}$ [$_{\text{CP}_3}$ wen$_1$ er t_2 gesagt hat] [$_{\text{CP}_2}$ t'_1 dass Fritz t_1 liebt]]

b. $[_{CP_2}$ t'_1 dass Fritz t_1 liebt] weiß ich t_3 nicht $[_{CP_3}$ $[_{CP_3}$ wen_1 er t_2 gesagt hat] t'_2]

Abstracting away from the indeterminacy of rule application in the last two steps, the movement theory adopted so far tells us that this derivation is the only one that is compatible with locality and cyclicity constraints. Thus, in order to systematically rule out sentences of the type in (32), we must hope that there is nevertheless something wrong with this derivation, and indeed, it turns out that it is excluded by the economy constraint Fewest Steps as involving a PUB effect. To see this, note that CP_2 is extraposed in step (35) and later topicalized in the last step in (36-b). Therefore, two distinct movement operations apply to one and the same item in the course of the derivation, and a PUB violation occurs. This violation is derivable from Fewest Steps. The reason is that there is a competing derivation that generates the same S-structure representation as the derivation in (33) – (36) (henceforth, D_1), but does so with fewer checking operations and therefore blocks D_1.

In such a blocking derivation, intermediate CP_2 extraposition does not take place. However, D_2 in (37) cannot be the derivation that we are looking for:

(37) a. *D-Structure*:
– weiß ich $[_{CP_3}$ – er $[_{CP_2}$ – dass Fritz wen_1 liebt] gesagt hat] nicht

b. *Wh-Movement of ‘wen$_1$’ to SpecC$_2$*:
– weiß ich $[_{CP_3}$ – er $[_{CP_2}$ wen_1 dass Fritz t_1 liebt] gesagt hat] nicht

c. *Wh-Movement of ‘wen$_1$’ to SpecC$_3$*:
– weiß ich $[_{CP_3}$ wen_1 er $[_{CP_2}$ t'_1 dass Fritz t_1 liebt] gesagt hat] nicht

d. *Extraposition of CP$_3$*:
– weiß ich t_3 nicht $[_{CP_3}$ wen_1 er $[_{CP_2}$ t'_1 dass Fritz t_1 liebt] gesagt hat]

e. *Topicalization of CP$_2$*:
$[_{CP_2}$ t'_1 dass Fritz t_1 liebt] weiß ich t_3 nicht $[_{CP_3}$ wen_1 er t_2 gesagt hat]

D_2 differs minimally from D_1 in that intermediate (hence, invisible) CP_2 extraposition fails to occur. Given that *wh*-movement in step (37-b) invariably instantiates an [extr] feature on C_2, this derivation does not converge – the strong [extr] feature remains unchecked at the point in the derivation where Spell-Out applies. However, recall from the notion of reference set arrived at in chapter 2 (repeated here as (38)) that only derivations that converge can compete with respect to transderivational economy constraints like Fewest Steps.

(38) *Reference Set*:
Two derivations D_1 and D_2 are in the same reference set iff they yield the same LF output and converge at LF and PF.

Thus, D_2 does not compete with D_1 and consequently cannot block it via Fewest Steps.

But now consider derivation D_3 in (39):

(39) a. *D-Structure*:
– weiß ich [$_{CP_3}$ – er [$_{CP_2}$ – dass Fritz wen$_1$ liebt] gesagt hat] nicht

b. *Wh-Movement of 'wen$_1$' to SpecC$_3$*:
– weiß ich [$_{CP_3}$ wen$_1$ er [$_{CP_2}$ – dass Fritz t$_1$ liebt] gesagt hat] nicht

c. *Extraposition of CP$_3$*:
– weiß ich t$_3$ nicht [$_{CP_3}$ wen$_1$ er [$_{CP_2}$ – dass Fritz t$_1$ liebt] gesagt hat]

d. *Topicalization of CP$_2$*:
[$_{CP_2}$ – dass Fritz t$_1$ liebt] weiß ich t$_3$ nicht [$_{CP_3}$ wen$_1$ er t$_2$ gesagt hat]

D$_3$ generates the same S-structure representation as D$_1$ and D$_2$. As in D$_2$, CP$_2$ extraposition does not take place. What is more, though, in this case, *wh*-movement skips over the intervening SpecC$_2$ position and goes directly to SpecC$_3$. Therefore, an [extr] feature does not have to show up on C$_2$ at any point in the derivation, and CP$_2$ extraposition can be left undone without leading to a non-converging derivation. Since D$_3$ involves fewer instances of feature checking than D$_1$ and since both derivations are members of one and the same reference set, it follows that D$_1$ is blocked by D$_3$. D$_3$, however, involves a highly non-local movement step: *wen*$_1$ moves from its in situ position directly into its ultimate target position, SpecC$_3$. Such a kind of movement induces a fatal violation of the Barriers Condition (the initial, hence undeletable, trace t$_1$ itself is the offending trace here, since an intermediate trace is not established) and therefore creates strong ungrammaticality even in the case of argument movement.

This, in a nutshell, is the explanation for the impossibility of sentences like (1)/(32): *Wh*-extraction via SpecC automatically triggers CP extraposition, and subsequent CP topicalization creates a PUB effect, reducible to the Fewest Steps condition. Under this view, what is wrong with (1)/(32) is not that an unbound intermediate trace occurs but, on the contrary, that no such trace may occur, because of economy.

Given that Derivational [extr] Instantiation obligatorily inserts [extr] features only on finite C nodes, and not on non-finite C nodes (which we have to assume to permit *wh*-extraction from infinitival CPs that show up to the left of V in German; cf. (20)), nothing blocks the topicalization of remnant infinitives that contain an intermediate *wh*-trace, as in (15). These examples are repeated here:

(40) a. ??[$_{CP}$ t$'_1$ PRO t$_1$ Zu lesen]$_3$ weiß ich nicht [$_{CP}$ was$_1$ der Fritz t$_3$ abgelehnt hat]
to read know I now what ART Fritz refused has

b. ??[$_{CP}$ t$'_1$ PRO t$_1$ Zu beanstanden]$_3$ weiß ich nicht [$_{CP}$ was$_1$ t$_3$ sich nicht gehört]
to object to know I not what is not proper

A derivation of, e.g., (40-a) may look as follows:

(41) a. *D-Structure*:
– weiß ich [$_{\text{CP}_2}$ – der Fritz [$_{\text{CP}_3}$ – PRO was$_1$ zu lesen] abgelehnt hat] nicht

b. *Wh-Movement of 'was$_1$' to SpecC$_3$*:
– weiß ich [$_{\text{CP}_2}$ – der Fritz [$_{\text{CP}_3}$ was$_1$ PRO t$_1$ zu lesen] abgelehnt hat] nicht

c. *Wh-Movement of 'was$_1$' to SpecC$_2$*:
– weiß ich [$_{\text{CP}_2}$ was$_1$ der Fritz [$_{\text{CP}_3}$ t$'_1$ PRO t$_1$ zu lesen] abgelehnt hat] nicht

d. *Topicalization of CP$_3$*:
[$_{\text{CP}_3}$ t$'_1$ PRO t$_1$ zu lesen] weiß ich [$_{\text{CP}_2}$ was$_1$ der Fritz t$_3$ abgelehnt hat] nicht

e. *Extraposition of CP$_2$*:
[$_{\text{CP}_3}$ t$'_1$ PRO t$_1$ zu lesen] weiß ich t$_2$ nicht [$_{\text{CP}_2}$ was$_1$ der Fritz t$_3$ abgelehnt hat]

This derivation involves a weak *wh*-island violation, which accounts for the intermediate status of the resulting sentence, but it does not involve a violation of the Fewest Steps condition (which would be much more severe). The reason for this is that *wh*-movement to SpecC$_3$ may take place without instantiating an [extr] feature on C$_3$ – Derivational [extr] Instantiation in (30) says something about finite clausal heads but not about non-finite ones.

3.3.2. A Consequence: PUB vs. Economy

At first sight, it seems that this approach faces a problem in view of examples like those in (42), which are well formed despite the fact that the *wh*-clauses do not show up in an extraposed position (cf. (27)):

(42) a. [$_{\text{CP}_2}$ Warum$_1$ C sie das t$_1$ getan hat] fragt sich Fritz t$_2$
why she$_{nom}$ that$_{acc}$ done has asks REFL Fritz$_{nom}$

b. [$_{\text{CP}_2}$ Wen$_1$ C sie t$_1$ liebt] ist mir t$_2$ leider unklar
whom$_{acc}$ she$_{nom}$ loves is me$_{dat}$ unfortunately unclear

Since *wh*-movement to SpecC instantiates an [extr] feature on the head of CP$_2$ according to (30), these examples must involve a combination of extraposition and subsequent topicalization of CP$_2$. Under a strict PUB account, we should therefore wrongly expect these examples to be ill formed. But given that PUB effects can be derived from the economy constraints Fewest Steps and Last Resort, this consequence does not arise. Recall that the derivation D$_1$ in (33) – (36) can only be blocked via Fewest Steps because there is a competing derivation in the same reference set (D$_3$) that involves one checking operation fewer. It follows that the examples in (42) are a problem for the present approach only if there is a competing derivation that involves fewer checking operations (and that violates some principle of grammar, like the Barriers Condition). However, there is no such derivation: The derivations of the examples in (42) each involve one instance of [+wh] feature checking, one instance of [extr] feature checking, and one instance of [top] feature checking; if any of these features remains unchecked, Last Resort is violated. The [+wh] feature is obligatory, the [extr] feature is obligatorily inserted after checking of the [+wh] feature, and the optional [top] feature is licensed by its creating a different LF representation (recall the theorem concerning optional movement features (120) of chapter 2). Thus, what differentiates between ill-formed examples like (32) and well-formed examples like those in (42) is the fact that the former can have a derivation in which an [extr] feature can be dispensed with (by skipping over an intermediate SpecC position in the course of *wh*-movement at the cost of a fatal locality violation), whereas the latter cannot have a derivation in which the [extr] feature is dispensed with (there is only one movement step of the *wh*-phrase in the first place).

Basically the same account can be given for the observation that clauses with internal *wh*-movement can show up not only in extraposed and topicalized positions, but also in scrambling positions. Consider the contrast in (43):

(43) a. *dass der Fritz überhaupt nicht [$_{CP_2}$ wie$_1$ du das Auto t$_1$ repariert
that ART Fritz$_{nom}$ absolutely not how you the car fixed
hast] wissen konnte
have know could

b. ??dass [$_{CP_2}$ wie$_1$ du das Auto t$_1$ repariert hast] der Fritz überhaupt
that how you the car fixed have ART Fritz$_{nom}$ absolutely
nicht t$_1$ wissen konnte
not know could

(43-b) has the same degree of grammaticality as regular instances of IP-internal finite [–wh] clauses in German. This means that it must not be ruled out by economy constraints. Indeed, given that (43-b) involves checking of the features [+wh], [extr] (generated by local *wh*-movement), and [scr] (optionally inserted on C), the derivation violates neither Fewest Steps nor Last Resort. Hence, (relative) wellformedness results despite the fact that the sequence of extraposition and scrambling applying to one and the same item (CP$_2$) would violate the PUB.

The ultimate rationale behind both cases discussed in this subsection is that the otherwise optional [extr] feature is turned into an obligatory feature in certain contexts: If the feature [extr] cannot be avoided, the derivation in which it occurs is legitimate even though another feature like [top] or [scr] is subsequently checked by the same item, in the same way that a derivation is legitimate in which an obligatory [Case] feature is checked prior to a [+wh] feature (see chapter 2).[13]

---

[13]This analysis also accounts for the fact that embedded V/2 clauses can undergo topicalization, as an alternative to extraposition; compare (31-a) with (i-a).

(i) a. [$_{CP}$ Den Peter$_2$ habe die Claudia t$_2$ geküßt]$_1$ sagte er t$_1$
ART Peter$_{acc}$ has$_{s}$ub ART Claudia$_{nom}$ kissed said he

4. A Problem and a Conclusion

In the final section of this chapter, I would like to first discuss one construction that at first sight supports the view that intermediate traces cannot occur unbound; and after that, I draw a conclusion.

4.1. Subextraction and Intermediate Traces

So far, I have discussed data that at first sight seem to suggest that unbound intermediate traces are not permitted, and I have shown that these data can and, in fact, must be accounted for without such a construction-specific assumption. It should be noted, however, that there are other constructions discussed by den Besten & Webelhuth (1990) and Fanselow (1992; 1993) that initially appear to support a prohibition against intermediate unbound traces. One particularly interesting case involves subextraction (*was für* split) in German; a standard example is given in (44):

(44) Was$_1$ hat er [$_{\text{NP}}$ t$_1$ für Bücher]$_2$ gelesen ?
what has he$_{nom}$ for books read

Here, the *wh*-item *was*$_1$ has been extracted out of NP$_2$, stranding NP$_2$ in situ. There seems to be no a priori reason why the remnant NP$_2$ should not be able to undergo long-distance topicalization across a *wh*-island created by subextraction itself. However, the result is strongly ungrammatical, as (45) shows:

(45) *[$_{\text{NP}}$ t$_1$ Für Bücher]$_2$ weiß ich nicht [$_{\text{CP}}$ was$_1$ er t$_2$ gelesen hat]
for books know I not what he read has

Why should this be so? The answer provided by den Besten & Webelhuth (1990, 81) is that subextraction as in (44) invariably involves an intermediate NP-internal trace, as in (46):[14]

(46) Was$_1$ ... [$_{\text{NP}}$ t$'_1$ [$_{\text{NP}}$ t$_1$ [$_{\text{PP}}$ für Bücher]]]$_2$ ...

On this approach, (45) has an unbound intermediate trace of *was*$_1$ in the fronted remnant NP, as shown in (47), and this intermediate trace, according to den Besten & Webelhuth, is responsible for the illformedness of the sentence.

---

b. ?*dass [$_{\text{CP}}$ den Peter$_2$ habe die Claudia t$_2$ geküßt]$_1$ noch keiner t$_1$ gesagt hat
that ART Peter$_{acc}$ has$_s$ub ART Claudia$_{nom}$ kissed yet no-one said has

However, what is left unexplained by this analysis is that embedded V/2 clauses still cannot undergo scrambling; (i-b) might be marginally better than (31-a), but it is far from perfect.

[14]The structure of NPs of the form *was für N* is a bit more involved under den Besten & Webelhuth's assumptions (e.g., the highest NP is analysed as a DP), but these details are not important in the present context.

(47) *[$_{\text{NP}}$ t'_1 [$_{\text{NP}_4}$ t_1 [$_{\text{PP}_3}$ Für Bücher]]]$_2$ weiß ich nicht [$_{\text{CP}}$ was$_1$ er t_2 gelesen hat]
for books know I not what he read has

Such a solution is no longer available under the present approach. But there is evidence that this approach cannot be maintained anyway and that there might be an independent factor at play in (45). To see this, note first that there is an alternative derivation of (47) that fails to produce an unbound intermediate trace. Given that *für Bücher* is an XP constituent (a PP, according to den Besten & Webelhuth's own assumptions), it is unclear what precludes it from undergoing long-distance topicalization out of NP$_2$; such a derivation would not create any unbound trace whatsoever:[15]

(48) *[$_{\text{PP}_3}$ Für Bücher] weiß ich nicht [$_{\text{CP}}$ was$_1$ er [$_{\text{NP}}$ t'_1 [$_{\text{NP}_4}$ t_1 t_3]]$_2$ gelesen hat]
for books know I not what he read has

Alternatively, what has undergone topicalization in (45) might be neither PP$_2$, as in (48), nor NP$_2$, as in (47), but rather the intermediate NP$_4$, as in (49):

(49) *[$_{\text{NP}_4}$ t_1 [$_{\text{PP}_3}$ Für Bücher]] weiß ich nicht [$_{\text{CP}}$ was$_1$ er [$_{\text{NP}}$ t'_1 t_4]$_2$ gelesen hat]
for books know I not what he read has

In this case, topicalization would have created an unbound trace of *was*, but, crucially, not an unbound intermediate trace, and hence the illformedness of the sentence would still be unexpected.

Next, consider the example in (50):

(50) *[$_{\text{NP}}$ Bücher]$_3$ weiß ich nicht [$_{\text{CP}}$ was$_1$ er [$_{\text{NP}}$ t_1 für t_3]$_2$ gelesen hat]
books know I not what he for read has

(50) is just as ungrammatical as (45), and we would like to be able to state that this is so for the very same reason. Indeed, it seems to be the case that normally neither the remnant NP that is created after subextraction (i.e., after movement of *was*) nor a part of this remnant NP can undergo any kind of movement on its own. However, there is one exception to this rule, and that is particularly interesting, given the approach to extraposition developed in chapter 4. It turns out that the sequence *für (D) N* can (marginally) undergo extraposition after subextraction to SpecC, as in (51):[16]

---

[15] Note that it does not seem possible to argue that NP$_2$ in this context is a barrier that blocks extraction of PP$_3$; NP$_2$ occupies a complement position; and, of course, it does not block subextraction of *was*$_1$. In addition, recall that the governing verb *lesen* ('read') belongs to the class of verbs that in principle permit extraction out of a direct object NP in situ, as a lexical property. – I hasten to add that for some unclear reason, subextraction of *was*, although sensitive to the position of the NP from which extraction takes place (cf. den Besten (1981), Grewendorf (1989), Diesing (1992), and many others), appears insensitive to lexical properties of the governing verb; cf. Müller (1995, 66) for discussion.

[16] This sentence is not completely perfect; but what is relevant in the present context is that it has the marked status that is often associated with heavy NP shift in German.

(51) a. ?Was$_1$ hat er t$_2$ gelesen [$_{NP}$ t$'_1$ [$_{NP}$ t$_1$ für Bücher]]$_2$?
what has he read for books
b. ?Was$_1$ hat er [$_{NP}$ t$'_1$ t$_4$]$_2$ gelesen [$_{NP_4}$ t$_1$ für Bücher] ?
what has he read for books
c. ?Was$_1$ hat er [$_{NP}$ (t$'_1$) [$_{NP}$ t$_1$ t$_3$]]$_2$ gelesen [$_{PP_3}$ für Bücher] ?
what has he read for books

As indicated, there are three analyses that are a priori possible. Either the extraposed string *für Bücher* is a bare PP without any unbound trace, as in (51-b), or it is a remnant NP, as in (51-a) and (51-b). In (51-a) NP$_2$ is extraposed, and an intermediate unbound trace of *was* is created; in contrast, in (51-b) NP$_4$ is extraposed, thereby producing only an initial unbound trace of *was*.[17] Given that extraposition in (51) is not completely well formed, and given that the extraposed string in (51) shows heaviness (or focussing) effects, the analysis in terms of NP extraposition, i.e., Heavy NP Shift, might indeed emerge as the correct one. But if this is so, we arrive at the result that the data in (51) might possibly suggest an analysis where a trace of *was* (perhaps even an intermediate trace of *was*) occurs unbound; in that case, even the evidence from subextraction does not warrant a condition like (5) that generally precludes unbound intermediate traces.

There are other striking peculiarities of the construction. Consider, e.g., the sentences in (52):

(52) a. Warum hat Karl der Maria [$_{NP}$ was$_1$ für Bücher] gegeben ?
why has Karl ART Maria$_{dat}$ what for books$_{acc}$ given
b. *Warum hat Karl was$_1$ der Maria [$_{NP}$ t$_1$ für Bücher] gegeben ?
why has Karl what ART Maria$_{dat}$ for books$_{acc}$ given

(52-a) shows that *wh*-NPs of the *was für N* kind can show up in situ in multiple questions. In the ungrammatical example (52-b), subextraction has taken place in this configuration. However, the relevant difference between this and well-formed cases of subextraction is that *was*$_1$ in (52-b) has not undergone *wh*-movement but scrambling. Now recall that *wh*-scrambling in German has been ruled out by invoking economy constraints (Fewest Steps and Last Resort) in chapter 2. Still, though, there is a clear contrast between illicit *wh*-scrambling, as in (53), and illicit subextraction to a scrambling position, as in (52-b); and this contrast suggests that there is an additional

[17] Note that the traces of *was* in the extraposed item must be unbound at S-structure under the approach to (Anti-) Freezing developed in chapter 4: A combination of extraction from XP and extraposition of XP is possible only if extraposition applies after extraction (because of the Barriers Condition) and ends up at least in the same minimal residue (because of the Strict Cycle Condition) – in this case, in the minimal residue of C.

violation of some constraint in (52-b).[18]

(53) ??Wie$_1$ hat [$_{IP}$ was$_2$ [$_{IP}$ der Fritz t$_1$ t$_2$ repariert]] ?
how has what$_{acc}$ ART Fritz fixed

What this clearly shows is that subextraction constructions not only prohibit movement of the type in (45), but behave in an unexpected way in other respects, too. I take it that the correct descriptive generalization about subextraction options in German is the following: In the structure

(54) ... [$_{NP_2}$ was$_1$ für (D) N] ...

the only possible movement operations are (i) movement of *was*$_1$ to SpecC; (ii) movement of the complete NP$_2$; and (iii) movement of the string *für (D) N* to a right-adjoined position. An explanation for these facts is obviously desirable. However, I cannot do such a thing here. My goal was to show that these peculiarities associated with subextraction are not related to the issue of remnant movement and that a specific prohibition against unbound intermediate traces does not even suffice to exclude the sentence in (45).

4.2. Conclusion

Let me summarize the main findings of this chapter. I considered ungrammatical cases of remnant movement where an intermediate trace of *wh*-extraction shows up unbound; cf. (1) and (2), both repeated here:

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

b. *[$_{VP}$ t$_2$ Gesagt [$_{CP_2}$ t$'_1$ dass Fritz t$_1$ liebt]]$_3$ weiß ich nicht [$_{CP}$
said that Fritz$_{nom}$ loves know I$_{nom}$ not
wen$_1$ er t$_3$ hat]
who$_{acc}$ he$_{nom}$ has

I argued that a construction-specific statement along the lines of den Besten & Webelhuth (1990), to the effect that intermediate traces can simply never occur unbound (see (5)), is neither conceptually attractive nor empirically adequate. The most important empirical counter-argument was that sentences involving unbound intermediate traces in fronted incoherent infinitives (or in fronted VPs that contain non-extraposed incoherent infinitives) turned out to be possible, in contrast to sentences like those in (55)

[18] I assigned (53) a star in chapter 2; cf. (19-c) of that chapter. That (53) is assigned two question marks here is not supposed to mean that the sentence has improved in the meantime; rather, it reflects the fact that grammaticality judgements are to be viewed as relative, not as absolute.

(cf. (15) and (18)). The basic idea then was that the illformedness of the examples in (55) does not have anything to do with the fact that an intermediate trace occurs unbound. Rather, it is due to restrictions on extraposition. The explanation I advanced is not homogeneous – (55-a) and (55-b) are ruled out for different reasons: On the one hand, (55-b) is impossible because extraposition of CP_2 ends up in a position that is too low to avoid a Freezing effect (if *wh*-extraction from a finite CP is to apply, CP must be extraposed to the next higher CP). (55-a), on the other hand, is impossible because the fact that *wh*-extraction from a finite CP has applied signals that CP must have undergone extraposition in the derivation; thus, extraposition of CP_2 feeds subsequent topicalization applying to the same item, and a PUB effect results, reducible to the economy condition Fewest Steps. Thus, neither of the sentences in (55) demands a construction-specific assumption about unbound traces; in fact, the assumptions that I made about extraposition were motivated by sentences like (56), where no unbound trace shows up:

(56) a. (Ich weiß nicht) wen$_1$ er t_2 gesagt hat [$_{CP}$ t'_1 dass Claudia t_1 geküßt hat]$_2$
I know not whom he said has that Claudia kissed has

b. *(Ich weiß nicht) wen$_1$ er [$_{CP}$ t'_1 dass Claudia t_1 geküßt hat]$_2$ gesagt hat
I know not whom he that Claudia kissed has said has

The wellformedness of (56-a) follows under the present system if we assume that extraposition may end up in a CP-adjoined position; and the illformedness of (56-b) follows if we assume that *wh*-movement to the minimal residue of a finite C node obligatorily creates an [extr] feature on C.

Chapter 7

Concluding Remarks

In the preceding chapters, I have examined various instances of remnant movement constructions resulting from topicalization, *wh*-movement, scrambling, left dislocation, and extraposition. I have tried to show that there are a number of properties of remnant movement in German that are surprising under a representational approach (viz., violations of the Proper Binding Condition, Anti-Freezing effects, movement type asymmetries, and an apparent prohibition against unbound intermediate traces), and that these properties can be explained by invoking independently established constraints in a derivational approach to movement theory, along the lines of Chomsky's (1991; 1993; 1995) minimalist program. The constraints that turned out to be particularly relevant are the Barriers Condition, Fewest Steps, Last Resort, the Strict Cycle Condition, and the Minimal Link Condition (MLC), viewed here as a part of Last Resort. This approach relies on the assumption that movement operations are triggered by features, and the features that I have focussed on are [+wh], [top], [scr], and [extr].

Thus, on the empirical side my basic aim has been to show that for an adequate account of the properties of remnant movement in German, a derivational approach to movement theory is preferable to a representational one. On a more general note, it has often been observed that representational analyses of syntactic phenomena can usually be rephrased in derivational terms (and vice versa) without too much difficulty, so most syntactic phenomena do not force a decision as to what the overall organization of a grammar (representational or derivational) should look like. Therefore, as soon as we find a phenomenon that strongly suggests preferring one approach over the other, this result may have important repercussions that go beyond the question of an empirically adequate account of the phenomenon itself. Accordingly, I would like to contend that, on the conceptual side, the analysis developed here should be viewed as an argument for a type of grammar in which all constraints are either derivational or transderivational (and constraints that can be viewed as filters on surface representations do not exist at all).

At this point, we may note that the type of grammar presupposed throughout this study, while derivational in nature, remains a hybrid form because it incorporates representational residues like the concept of chain and, indeed, the notion of trace. As has been argued by Koster (1987), Brody (1995), and several others, this type of hybrid derivational grammar is characterized by a fundamental redundancy. Given the presence of traces, most instances of movement are encoded twice: by the movement operation itself, and by chain formation algorithms that link traces to their antecedents. I think there is a grain of truth in this critique, and, to end this study, I would like to briefly speculate on how this dilemma can be solved. As a starting point, it is worth noting that (assuming the MLC-based approach to Unambiguous Domination effects) none of the syntactic constraints that are crucial for the analyses given in the preceding chapters actually mention either traces or chains. In the present framework, these concepts are necessary mainly for semantic reasons, e.g., to provide semantic variable binding in the case of *wh*-operators, or to help locating lexical expressions in the positions where they are needed to obtain a compositional semantics at the level of LF (see especially the remarks on the copy theory in chapter 3). Thus, suppose that we were to give up the assumption that movement leaves a trace. This would leave most (if not all, given some obvious adjustments) of the syntactic reasonings in the previous chapters intact. However, what would have to be dropped is the idea that representations at the level of LF can be compositionally interpreted as such – i.e., what would be called for is a radically different concept of semantic interpretation. A possible framework would be one in which semantic interpretation and syntactic derivation go hand in hand, i.e., interpretation works derivationally, too; a semantic theory that is designed to meet this requirement is sketched in Sternefeld (1996). Of course, if this view is adopted, numerous questions and problems arise which I cannot pursue in this study. However, if this radical perspective on the syntax/semantics interface ultimately proves viable, we can conclude that the main conceptual argument of this monograph does not only support a hybrid derivational approach to grammar, but, in fact, a strictly derivational one.

References

Ackema, Peter, & Ad Neeleman. 1995. Optimal Questions. Ms., University of Utrecht.

Aoun, Joseph, Norbert Hornstein, & Dominique Sportiche. 1981. On Some Aspects of Wide Scope Interpretation. *Journal of Linguistic Research* 1:69-95.

Baker, Carl L. 1970. Notes on the Description of English Questions: the Role of an Abstract Question Morpheme. *Foundations of Language* 6:197-219.

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

Baltin, Mark. 1983. Extraposition: Bounding versus Government-Binding. *Linguistic Inquiry* 14:155-162.

Baltin, Mark. 1984. Extraposition Rules and Discontinuous Constituents. *Linguistic Inquiry* 15:157-163.

Baltin, Mark, & Paul Postal. 1996. More on Reanalysis Hypotheses. *Linguistic Inquiry* 27:127-145.

Barss, Andrew. 1984. Chain Binding. Ms., MIT, Cambridge, Mass.

Barss, Andrew. 1985. Adjunction and Reflexive Hierarchical Relations. Ms., MIT, Cambridge, Mass.

Barss, Andrew. 1986. *Chains and Anaphoric Dependence*. Doctoral dissertation, MIT, Cambridge, Mass.

Barss, Andrew, & Howard Lasnik. 1986. A Note on Anaphora and Double Objects. *Linguistic Inquiry* 17:347-354.

Bayer, Josef. 1990. *Directionality of Government and Logical Form: A Study of Focusing Particles and Wh-Scope*. Habilitation thesis, Universität Konstanz.

Bayer, Josef. 1990a. 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. 1994. Barriers for German. Ms., Universität Stuttgart.

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

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

Beck, Sigrid. 1996. Negative Islands and Reconstruction. In *On Extraction and Extraposition in German*, ed. Uli Lutz & Jürgen Pafel, 121-143. Amsterdam: Benjamins.

Beck, Sigrid, & Stephen Berman. 1996. Wh-Scope Marking: Direct vs. Indirect Dependency. In *Papers on Wh-Scope Marking*, ed. Uli Lutz & Gereon Müller, 59-83. Tübingen: Arbeitspapiere des SFB 340 no. 76.

Behagel, Otto. 1932. *Deutsche Syntax. Eine geschichtliche Darstellung. Band IV: Wortstellung. Periodenbau*. Heidelberg: Carl Winters Universitätsbuchhandlung.

Bennis, Hans, & Teun Hoekstra. 1984. Gaps and Parasitic Gaps. *The Linguistic Review* 4:29-87.

Besten, Hans den. 1981. Government, syntaktische Struktur und Kasus. In *Sprache: Formen und Strukturen*, ed. Manfred Kohrt & Jürgen Lenerz, 97-107. Tübingen: Niemeyer.
Besten, Hans den. 1989. *Studies in West Germanic Syntax*. Doctoral dissertation, University of Amsterdam.
Besten, Hans den & Jan Rutten. 1989. On Verb Raising, Extraposition and Free Word Order in Dutch. In *Sentential Complementation and the Lexicon*, ed. D. Jaspers et al., 41-56. Dordrecht: Foris.
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.
Bierwisch, Manfred. 1963. *Grammatik des deutschen Verbs*. Berlin: Akademieverlag.
Bobaljik, Jonathan, & Samuel Brown. 1997. Interarboreal Operations: Head Movement and the Extension Requirement. *Linguistic Inquiry* 28:345-356.
Borsley, Robert. 1990. A GPSG Approach to Breton Word Order. In *The Syntax of the Modern Celtic Languages* (Syntax and Semantics, vol. 23), ed. Randall Hendrick, 81-95. San Diego etc.: Academic Press.
Brandner, Eleonore. 1994. *Specifiers and Functional Projection*. Doctoral dissertation, Universität Stuttgart.
Branigan, Phil. 1992. *Subjects and Complementizers*. Doctoral dissertation, MIT, Cambridge, Mass.
Brody, Michael. 1995. *Lexico-Logical Form*. Cambridge, Mass.: MIT Press.
Browning, M.A. 1987. *Null Operator Constructions*. Doctoral dissertation, MIT, Cambridge, Mass.
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. 1993. Es gibt keine freie Wortstellung im Deutschen. Paper presented at the annual meeting of the DGfS, Jena.
Büring, Daniel. 1995. On the Base Position of Embedded Clauses in German. *Linguistische Berichte* 159:370-380.
Büring, Daniel. 1996. *The 59th Street Bridge Accent*. Doctoral dissertation, Universität Tübingen.
Büring, Daniel, & Katharina Hartmann. 1994. The Dark Side of Wh-Movement. *Linguistische Berichte* 149:56-74.
Büring, Daniel, & Katharina Hartmann. 1997. Doing the Right Thing. *The Linguistic Review* 14:1-42.
Cardinaletti, Anna. 1986. Topicalization in German: Movement to COMP or Base-Generation in TOP? *Groninger Arbeiten zur Germanistischen Linguistik* 28:202-231.
Cardinaletti, Anna, & Ian Roberts. 1991. Clause Structure and X-Second. Ms., University of Venice and Université de Genève. To appear in *Levels, Principles, and Processes: The Structure of Grammatical Representations*, ed. Wynn Chao & Geoffrey Horrock. Berlin: Mouton de Gruyter.
Cardinaletti, Anna & Michal Starke. 1995. Deficient Pronouns. Ms., University of Venice.
Chomsky, Noam. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton.
Chomsky, Noam. 1973. Conditions on Transformations. In *A Festschrift for Morris Halle*, ed. Stephen Anderson & Paul Kiparsky, 232-286. New York: Academic Press.
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. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1991. Some Notes on Economy of Derivation and Representation. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin, 417-454. 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 & Howard Lasnik. 1993. Principles and Parameters Theory. In *Syntax, vol. I*, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, & Theo Vennemann, 506-569. Berlin: de Gruyter.
Cinque, Guglielmo. 1977. The Movement Nature of Left Dislocation. *Linguistic Inquiry* 8:397-412.
Cinque, Guglielmo. 1990. *Types of A'-Dependencies*. Cambridge, Mass.: MIT Press.
Cole, Peter. 1982. Subjacency and Successive Cyclicity: Evidence from Ancash Quechua. *Journal of Linguistic Research* 2:35-58.
Collins, Chris. 1993. *Topics in Ewe Syntax*. Doctoral dissertation, MIT, Cambridge, Mass.
Collins, Chris. 1994. Economy of Derivation and the Generalized Proper Binding Condition. *Linguistic Inquiry* 25:45-61.
Collins, Chris. 1994a. Merge and Greed. Ms., Cornell University.
Collins, Chris. 1997. *Local Economy*. Cambridge, Mass.: MIT Press.
Costa, João. 1995. The Problems of Unambiguous Binding, Greed and Fewest Steps. Ms., Leiden University.
Culicover, Peter. 1991. Topicalization, Inversion, and Complementizers in English. Ms., The Ohio State University, Columbus.
Culicover, Peter, & Michael Rochemont. 1990. Extraposition and the Complement Principle. *Linguistic Inquiry* 21:23-47.
d'Avis, Franz-Josef. 1993. Zur Lage auf den W-Inseln. In *Extraktion im Deutschen 1*, ed. Franz-Josef d'Avis et al., 63-98. Stuttgart & Tübingen: Arbeitspapiere des SFB 340 no. 34.
d'Avis, Franz-Josef. 1996. On wh-Islands in German. In *On Extraction and Extraposition in German*, ed. Uli Lutz & Jürgen Pafel, 89-120. Amsterdam: Benjamins.
Dayal, Veneeta Srivastav. 1994. Scope Marking as Indirect Wh Dependency. *Natural Language Semantics* 2:137-170.
Dekydtspotter, Laurent. 1992. The Syntax of Predicate Clefts. *NELS* 22:119-133.
Diesing, Molly. 1992. *Indefinites*. Cambridge, Mass.: MIT Press.
Emonds, Joseph. 1976. *A Transformational Approach to English Syntax*. New York: Academic Press.
Emonds, Joseph. 1978. The Verbal Complex V'-V in French. *Linguistic Inquiry* 9:151-175.
Engdahl, Elisabet. 1986. *Constituent Questions*. Dordrecht: Reidel.
Epstein, Samuel David. 1992. Derivational Constraints on $\overline{\text{A}}$-Chain Formation. *Linguistic Inquiry* 23:235-259.
Epstein, Samuel David. 1995. Un-Principled Syntax and the Derivation of Syntactic Relations. Ms., Harvard University.
Fanselow, Gisbert. 1983. *Zu einigen Problemen von Kasus, Rektion und Bindung in der deutschen Syntax*. M.A. Thesis, Universität Konstanz.

Fanselow, Gisbert. 1987. *Konfigurationalität*. Tübingen: Narr.
Fanselow, Gisbert. 1988. Aufspaltung von NPn und das Problem der 'freien' Wortstellung. *Linguistische Berichte* 114:91-113.
Fanselow, Gisbert. 1989. Coherent Infinitives in German. In *Syntactic Phrase Structure Phenomena in Noun Phrases and Sentences*, ed. Christa Bhatt, Elisabeth Löbel, & Claudia Schmidt, 1-16. Amsterdam: Benjamins.
Fanselow, Gisbert. 1990. Scrambling as NP-movement. In *Scrambling and Barriers*, ed. Günther Grewendorf & Wolfgang Sternefeld, 113-140. Amsterdam: Benjamins.
Fanselow, Gisbert. 1991. *Minimale Syntax*. Habilitation thesis, Universität Passau.
Fanselow, Gisbert. 1992. Deplazierte Argumente. Ms., Universität Stuttgart.
Fanselow, Gisbert. 1993. The Return of the Base Generators. Ms., Universität Passau. (Also in *Groninger Arbeiten zur Germanistischen Linguistik* 36:1-74.)
Fanselow, Gisbert. 1997. Features, Θ-Roles, and Free Constituent Order. Ms., Universität Potsdam.
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.
Ferguson, Scott, & Erich Groat. 1994. Defining "Shortest Move". Ms., Harvard University.
Fiengo, Robert. 1977. On Trace Theory. *Linguistic Inquiry* 8:35-61.
Fox, Danny. 1994. Quantifer Scope in VP Ellipsis. Ms., MIT, Cambridge, Mass.
Fox, Danny. 1995. Letter to the author.
Frampton, John. 1990. Parasitic Gaps and the Theory of Wh-Chains. *Linguistic Inquiry* 21:49-77.
Frampton, John, & Sam Gutmann. 1997. Eliminating Comparison of Derivations. Ms., Northeastern University.
Frank, Robert, Young-Suk Lee, & Owen Rambow. 1992. Scrambling as Non-Operator Movement and the Special Status of Subjects. In *Proceedings of the Third Leiden Conference for Junior Linguists*, ed. Sjef Barbiers, Marcel den Dikken, & Claartje Levelt, 135-154. Leiden: University of Leiden.
Freidin, Robert. 1978. Cyclicity and the Theory of Grammar. *Linguistic Inquiry* 9:519-549.
Freidin, Robert. 1992. *Foundations of Generative Syntax*. Cambridge, Mass.: MIT Press.
Frey, Werner, & Thilo Tappe. 1991. Zur Interpretation der X-bar-Theorie und zur Syntax des Mittelfeldes. Grundlagen eines GB-Fragmentes. Ms., Universität Stuttgart.
Fukui, Naoki. 1993. Parameters and Optionality. *Linguistic Inquiry* 24:399-420.
Fukui, Naoki. 1993a. A Note on Improper Movement. *The Linguistic Review* 10:111-126.
Gärtner, Hans-Martin, & Markus Steinbach. 1994. Economy, Verb Second, and the SVO-SOV Distinction. *Working Papers in Scandinavian Syntax* 53:1-59.
Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, & Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwell.
Geilfuß, Jochen. 1988. Preposition Stranding, Barrieren und Relativierte Minimalität. Ms., Universität Tübingen.
Georgopoulos, Carol. 1991. *Syntactic Variables*. Dordrecht: Kluwer.
Giusti, Giuliana. 1986. On the Lack of Wh-Infinitives with 'zu' and the Projection of COMP in German. *Groninger Arbeiten zur Germanistischen Linguistik* 28:115-169.
Grewendorf, Günther. 1987. Kohärenz und Restrukturierung. Zu verbalen Komplexen im Deutschen. In *Neuere Forschungen zur Wortbildung und Historiographie der Linguistik*, ed. Brigitte Asbach-Schnitger & Johannes Roggenhofer, 123-144. Tübingen: Narr.
Grewendorf, Günther. 1988. *Aspekte der deutschen Syntax*. Tübingen: Narr.
Grewendorf, Günther. 1989. *Ergativity in German*. Dordrecht: Foris.

Grewendorf, Günther. 1990. Verb-Bewegung und Negation im Deutschen. Ms. Universität Frankfurt/Main.

Grewendorf, Günther. 1994. Variable Binding and Reconstruction. Ms., Universität Frankfurt/Main. (*Sprachwissenschaft in Frankfurt*, Report no. 12.)

Grewendorf, Günther. 1995. German – A Grammatical Sketch. In *Syntax, vol. II*, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, & Theo Vennemann, 1288-1319. Berlin: de Gruyter.

Grewendorf, Günther, & Joachim Sabel. 1994. Long Scrambling and Incorporation. *Linguistic Inquiry* 25:263-308.

Grimshaw, Jane. 1994. Projection, Heads, and Optimality. Ms., Rutgers University. To appear in *Linguistic Inquiry*.

Guéron, Jacqueline, & Robert May. 1984. Extraposition and Logical Form. *Linguistic Inquiry* 15:1-31.

Haider, Hubert. 1983. Connectedness Effects in German. *Groninger Arbeiten zur Germanistischen Linguistik* 23:82-119.

Haider, Hubert. 1986. Affect α: A Reply to Lasnik and Saito,'On the Nature of Proper Government'. *Linguistic Inquiry* 17:113-126.

Haider, Hubert. 1988. Matching Projections. In *Constituent Structure*, ed. Anna Cardinaletti, Guglielmo Cinque, & Giuliana Giusti, 101-123. Dordrecht: Foris.

Haider, Hubert. 1990. Topicalization and Other Puzzles of German Syntax. In *Scrambling and Barriers*, ed. Günther Grewendorf & Wolfgang Sternefeld, 93-112. Amsterdam: Benjamins.

Haider, Hubert. 1991, Fakultativ kohärente Infinitivkonstruktionen im Deutschen. Ms., Universität Stuttgart. (Arbeitspapiere des SFB 340 no. 17.)

Haider, Hubert. 1993. *Deutsche Syntax – Generativ*. Tübingen: Narr.

Haider, Hubert. 1994. Detached Clauses – The Later the Deeper. Ms., Universität Stuttgart. (Arbeitspapiere des SFB 340 no. 41.)

Haider, Hubert. 1995. Extraposition. Ms., Universität Stuttgart.

Heim, Irene. 1982, *The Semantics of Definite and Indefinite Noun Phrases*. Doctoral dissertation, University of Massachusetts, Amherst.

Heim, Irene. 1987. Where Does the Definiteness Restriction Apply? In *The Representation of (In)definiteness*, ed. Eric Reuland & Alice ter Meulen, 21-42. Cambridge, Mass.: MIT Press.

Heim, Irene. 1989. *Survey of Formal Semantics*. Ms., MIT, Cambridge, Mass.

Heim, Irene. 1992. Remarks on Wh-Questions. Ms., MIT, Cambridge, Mass.

Hendrick, Randall. 1990. Breton Pronominals, Binding, and Barriers. In *The Syntax of the Modern Celtic Languages* (Syntax and Semantics, vol. 23), ed. Randall Hendrick, 121-165. San Diego etc.: Academic Press.

Henscheid, Eckhard. 1983. *Wie Max Horkheimer einmal sogar Adorno hereinlegte*. Zürich: Haffmanns.

Heycock, Caroline. 1995. Asymmetries in Reconstruction. *Linguistic Inquiry* 26:547-570.

Heycock, Caroline, & Anthony Kroch. 1997. Pseudocleft Connectivity: Implications for the LF Interface Level. Ms., University of Edinburgh and University of Pennsylvania.

Hinrichs, Erhard, & Tsuneko Nakazawa. 1994. Partial VP and Split NP Topicalization in German: An HPSG Analysis. Ms., Universität Tübingen. (Arbeitspapiere des SFB 340, no. 58.)

Höhle, Tilman. 1990. Die w...w-Konstruktion im Deutschen. Ms., Universität Tübingen.

Höhle, Tilman. 1991. On Reconstruction and Coordination. In *Representation and Derivation in*

the Theory of Grammar, ed. Hubert Haider & Klaus Netter, 139-197. Dordrecht: Kluwer.
Höhle, Tilman. 1996. The W-...W-Construction: Appositive or Scope Indicating. In *Papers on Wh-Scope Marking*, ed. Uli Lutz & Gereon Müller, 37-58. Tübingen: Arbeitspapiere des SFB 340 no. 76.
Hoekstra, Teun. 1983. The Distribution of Sentential Complements. In *Linguistics in the Netherlands 1983*, ed. Hans Bennis & W.U.S. van Lessen Kloeke. Dordrecht: Foris.
Hoekstra, Teun. 1984. *Transitivity*. Dordrecht: Foris.
Holmberg, Anders. 1986. *Word Order and Syntactic Features in the Scandinavian Languages and English*. Stockholm: Department of General Linguistics.
Holmberg, Anders. 1997. Object Shift and Phonological Features. Ms., University of Tromsø.
Hoof, Hanneke van. 1995. Left Dislocation and Split Topics in Brabant Dutch. Ms., Universität Tübingen.
Hoop, Helen de. 1992. *Case Configuration and Noun Phrase Interpretation*. Doctoral dissertation, Rijksuniversiteit Groningen.
Huang, Cheng-Teh James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Doctoral dissertation, MIT, Cambridge, Mass.
Huang, Cheng-Teh James. 1993. Reconstruction and the Structure of VP: Some Theoretical Consequences. *Linguistic Inquiry* 24:103-138.
Jackendoff, Ray. 1990. On Larson's Account of the Double Object Construction. *Linguistic Inquiry* 21:427-454.
Jacobs, Joachim. 1982. *Syntax und Semantik der Negation im Deutschen*. München: Fink.
Johnson, Kyle. 1985. *A Case for Movement*. Doctoral dissertation, MIT, Cambridge, Mass.
Johnson, Kyle. 1991. Object Positions. *Natural Language and Linguistic Theory* 9:577-636.
Kang, Jung-Goo, & Gereon Müller. 1994. Kopiertheorie gegen Rekonstruktion. *Linguistische Berichte* 152:261-280
Karttunen, Lauri. 1977. Syntax and Semantics of Questions. *Linguistics and Philosophy* 1:3-44.
Kayne, Richard. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.
Kayne, Richard. 1985. Principles of Particle Constructions. In *Grammatical Representation*, ed. Jacqueline Guéron, Hans-Georg Obenauer, & Jean-Yves Pollock, 101-140. Dordrecht: Foris.
Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
Kim, Shin-Sook. 1996. Bindungsrekonstruktion vs. Skopusrekonstruktion. *Linguistische Berichte* 165:401-431.
Kitagawa, Yoshihisha. 1986. *Subjects in Japanese and English*. Doctoral dissertation, University of Massachusetts, Amherst.
Kitahara, Hisatsugu. 1993. Deducing 'Superiority' Effects from the Shortest Chain Requirement. *Harvard Working Papers in Linguistics* 3:109-119.
Kitahara, Hisatsugu. 1994. Restricting Ambiguous Rule-Application. In *Formal Approaches to Japanese Linguistics 1*, ed. Masatoshi Koizumi & Hiroyuki Ura, 179-209. Cambridge, Mass.: MITWPL.
Kitahara, Hisatsugu. 1994a. A Minimalist Analysis of Cross-Linguistically Variant CED Phenomena. *NELS* 24:241-253.
Kitahara, Hisatsugu. 1995. Target α. *Linguistic Inquiry* 26:47-77.
Kohrt, Manfred. 1974. 'Command' and Rightward Movement Rules. *Linguistische Berichte* 30:26-32.
Koizumi, Masatoshi. 1995, *Phrase Structure in Minimalist Syntax*. Doctoral dissertation, MIT, Cambridge, Mass.
Kolb, Hans-Peter. 1995. Categories, Segments, and Other Dangerous Things. Ms., Universität Tübingen.

Koopman, Hilda, & Dominique Sportiche. 1991. The Position of Subjects. *Lingua* 85:211-258.

Koster, Jan. 1978. *Locality Principles in Syntax*. Dordrecht: Foris.

Koster, Jan. 1986. The Relation between pro-drop, Scrambling, and Verb Movements. Ms., Rijksuniversiteit Groningen.

Koster, Jan. 1987. *Domains and Dynasties*. Dordrecht: Foris.

Kratzer, Angelika. 1991. The Representation of Focus. In *Semantics. An International Handbook of Contemporary Research*, ed. Arnim von Stechow & Dieter Wunderlich, 825-834. Berlin: de Gruyter.

Larson, Richard. 1988. On the Double Object Construction. *Linguistic Inquiry* 19:335-391.

Larson, Richard. 1990. Double Objects Revisited. *Linguistic Inquiry* 21:589-632

Larson, Richard. 1991. *Promise* and the Theory of Control. *Linguistic Inquiry* 22:103-139.

Lasnik, Howard. 1993. *Lectures on Minimalist Syntax*. University of Connecticut Working Papers in Linguistics 1, MIT, Cambridge, Mass.

Lasnik, Howard. 1995. Case and Expletives Revisited: On Greed and Other Human Failings. *Linguistic Inquiry* 26:615-633.

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

Lasnik, Howard, & Mamoru Saito. 1992. *Move* α. Cambridge, Mass.: MIT Press.

Lasnik, Howard, & Juan Uriagereka. 1988. *A Course in GB Syntax*. Cambridge, Mass.: MIT Press.

Lenerz, Jürgen. 1995. Klammerkonstruktionen. In *Syntax, vol. II*, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, & Theo Vennemann, 1266-1276. Berlin: de Gruyter.

Longobardi, Giuseppe. 1985. Connectedness, Scope, and C-command. *Linguistic Inquiry* 16:162-193.

Lutz, Uli, & Susanne Trissler. 1992. Einige Ęberlegungen zu syntaktischen +w-Merkmalen, Interrogativsätzen und w-Phrasen im Deutschen. Ms. Universität Tübingen.

Lutz, Uli, & Gereon Müller (eds). 1996. *Papers on Wh-Scope Marking*. Tübingen: Arbeitspapiere des SFB 340 no. 76.

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

McCloskey, James. 1990. Resumptive Pronouns, A-bar Binding, and Levels of Representation in Irish. In *The Syntax of the Modern Celtic Languages* (Syntax and Semantics, vol. 23), ed. Randall Hendrick, 199-248. San Diego: Academic Press.

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.

Mahajan, Anoop. 1992. The Specificity Condition and the CED. *Linguistic Inquiry* 23:510-516.

Mahajan, Anoop. 1994. Two Apparent Rightward Movement Rules in Hindi and Their Implications. Ms., UCLA.

Manzini, Rita. 1992. *Locality*. Cambridge, Mass: MIT Press.

Marácz, László. *Asymmetries in Hungarian*. Doctoral dissertation, Rijksuniversiteit Groningen.

Marantz, Alec. 1995. The Minimalist Program. In *Government and Binding Theory and the Minimalist Program*, ed. Gert Webelhuth, 349-382. Oxford: Blackwell.

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

May, Robert. 1979. Must COMP-to-COMP Movement Be Stipulated? *Linguistic Inquiry* 10:719-725.

May, Robert. 1985. *Logical Form*. Cambridge, Mass.: MIT Press.

Meinunger, André. 1995. *Discourse Dependent DP (De-)Placement*. Doctoral dissertation. Universität Potsdam.
Müller, Gereon. 1991. Abstrakte Inkorporation. In *DET, COMP, INFL*, ed. Susan Olsen & Gisbert Fanselow, 155-202. Tübingen: Niemeyer.
Müller, Gereon. 1993. *On Deriving Movement Type Asymmetries*. Doctoral dissertation. Universität Tübingen.
Müller, Gereon. 1995. *A-bar Syntax*. Berlin: Mouton de Gruyter.
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, & Wolfgang Sternefeld. 1990. Improper Movement. Ms., Universität Konstanz.
Müller, Gereon, & Wolfgang Sternefeld. 1993. Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24:461-507.
Müller, Gereon, & Wolfgang Sternefeld. 1994. Scrambling as A-bar Movement. In *Studies on Scrambling*, ed. Norbert Corver & Henk van Riemsdijk, 331-385. Berlin: Mouton de Gruyter.
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.
Neeleman, Ad. 1994. Scrambling as a D-structure Phenomenon. In *Studies on Scrambling*, ed. Norbert Corver & Henk van Riemsdijk, 387-429. Berlin: Mouton de Gruyter.
Nerbonne, John. 1994. Partial Verb Phrases and Spurious Ambiguities. In *German in Head-Driven Phrase Structure Grammar*, ed. John Nerbonne, Klaus Netter, & Carl Pollard, 109-150. Stanford: CSLI Lecture Notes.
Ormazabal, Javier, Juan Uriagereka, & Maria Uribe-Etxebarria. 1994. Word Order and Wh-Movement: Towards a Parametric Account. Ms., University of Connecticut & University of Maryland.
Paul, Hermann. 1919. *Deutsche Grammatik. Band III, Teil IV: Syntax (erste Hälfte)*. Halle: Niemeyer.
Perlmutter, David, & Scott Soames. 1979. *Syntactic Argumentation and the Structure of English*. Berkeley: The University of California Press.
Pesetsky, David. 1982. *Paths and Categories*. Doctoral dissertation, MIT, Cambridge, Mass.
Pesetsky, David. 1987. Wh-in-Situ: Movement and Unselective Binding. In *The Representation of (In)Definiteness*, ed. Eric Reuland & Alice ter Meulen, 98-129. Cambridge, Mass.: MIT Press.
Plessis, Hans du. 1977. Wh-Movement in Afrikaans. *Linguistic Inquiry* 8:723-726.
Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20:365-424.
Postal, Paul. 1972. On Some Rules That Are Not Successive Cyclic. *Linguistic Inquiry* 3:211-222.
Pullum, Geoffrey. 1979. *Rule Interaction and the Organization of a Grammar*. New York: Garland.
Pullum, Geoffrey. 1992. The Origins of the Cyclic Principle. *Proceedings of the Regional Meeting of the Chicago Linguistic Society* 28.2:209-236.
Reinhart, Tanya. 1983. *Anaphora and Semantic Interpretation*. London: Croom Helm.

Reinhart, Tanya. 1994. Wh-in situ in the Framework of the Minimalist Program. Ms., University of Utrecht.

Reis, Marga. 1996. Extraction from V/2 Clauses in German? In *On Extraction and Extraposition in German*, ed. Uli Lutz & Jürgen Pafel, 45-88. Amsterdam: Benjamins.

Reis, Marga, & Inger Rosengren. 1992. What Do Wh-Imperatives Tell Us About Wh-Movement? *Natural Language and Linguistic Theory* 10:79-118.

Reuland, Eric, & Wim Kosmeijer. 1993. Projecting Inflected Verbs. In *The Parametrization of Universal Grammar*, ed. Gisbert Fanselow, 37-71. Amsterdam: Benjamins.

Riemsdijk, Henk van. 1978. *A Case Study in Syntactic Markedness*. 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.

Riemsdijk, Henk van. 1989. Movement and Regeneration. In *Dialectal Variation and the Theory of Grammar*, ed. Paola Benincá, 105-136. Dordrecht: Foris.

Rivero, Maria Luisa. 1991. Long Head Movement and Negation: Serbo-Croatian vs. Slovak and Czech. *The Linguistic Review* 8:319-351.

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

Rizzi, Luigi. 1996. Residual Verb Second and the *Wh*-Criterion. In *Parameters and Functional Heads*, ed. Adriana Belletti & Luigi Rizzi, 63-90. New York & Oxford: Oxford University Press.

Roberts, Ian. 1990. Some Notes on VP-Fronting and Head Government. In *Grammar in Progress. Glow Essays for Henk van Riemsdijk*, ed. Joan Mascaro and Marina Nespor, 387-396. Dordrecht: Foris.

Roberts, Ian. 1991. Excorporation and Minimality. *Linguistic Inquiry* 22:209-218.

Roberts, Ian. 1993. *Verbs and Diachronic Syntax*. Dordrecht: Kluwer.

Rochemont, Michael. 1978. *A Theory of Stylistic Rules in English*. Doctoral dissertation, University of Massachusetts, Amherst.

Rochemont, Michael. 1989. Topic Islands and the Subjacency Parameter. *Canadian Journal of Linguistics* 34:145-170.

Rochemont, Michael, & Peter Culicover. 1991. *English Focus Constructions and the Theory of Grammar*. Cambridge: Cambridge University Press.

Rohrbacher, Bernhard. 1994. *The Germanic Languages and the Full Paradigm: A Theory of V to I Raising*. Doctoral dissertation, University of Massachusetts, Amherst.

Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Doctoral dissertation, MIT, Cambridge, Mass. Appeared 1986 as *Infinite Syntax*. Norwood, New Jersey: Ablex Publishing Corporation.

Rudin, Catherine. 1988. On Multiple Questions and Multiple WH Fronting. *Natural Language and Linguistic Theory* 6:445-501.

Sabel, Joachim. 1993. W-Infinitive, Nominativsubjekte und das Wh-Criterion. Ms., Universität Frankfurt. (*Sprachwissenschaft in Frankfurt*, Report no. 11.)

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

Sabel, Joachim. 1995a. Intermediate Traces, Adjunction Movement, and Minimize Chain Links. Ms., Universität Frankfurt/Main.

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.

Saddy, Douglas. 1991. Wh-Scope Mechanisms in Bahasa Indonesia. In *More Papers on Wh-Movement*, ed. Lisa Cheng & Hamida Demirdash, 182-218. MIT, Cambridge, Mass.: MIT-WPL 15.

Saito, Mamoru. 1985. *Some Asymmetries in Japanese and their Theoretical Implications*. Doctoral dissertation, MIT, Cambridge, Mass.

Saito, Mamoru. 1989. Scrambling as Semantically Vacuous A-bar Movement. In *Alternative Conceptions of Phrase Structure*, ed. Mark Baltin & Anthony Kroch, 182-200. Chicago: The University of Chicago Press.

Saito, Mamoru. 1992. Long Distance Scrambling in Japanese. *Journal of East Asian Linguistics* 1:69-118.

Sauerland, Uli. 1996. Scrambling and Interpretability. Ms., MIT, Cambridge, Mass.

Schafer, Robin. 1995. Nonfinite Predicate-Initial Constructions in Breton. *Glot International* 1/8:12-13.

Schmidt, Clauda Maria. 1992. Eine mediale Klitikposition. Ms., Universität zu Köln.

Schwartz, Bonnie, & Sten Vikner. 1996. The Verb Always Leaves IP in V2 Clauses. In *Parameters and Functional Heads*, ed. Adriana Belletti & Luigi Rizzi, 11-62. New York & Oxford: Oxford University Press.

Speas, Margaret. 1990. *Phrase Structure in Natural Language*. Dordrecht: Kluwer.

Sportiche, Dominique. 1988. A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry* 19:33-60.

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

Stechow, Arnim von. 1984. Gunnar Bech's Government and Binding Theory. *Linguistics* 22:225-241.

Stechow, Arnim von. 1992. Kompositionsprinzipien und grammatische Struktur. In *Biologische und soziale Grundlagen der Sprache*, ed. Peter Suchsland, 175-248. Tübingen: Niemeyer.

Stechow, Arnim von. 1993. Die Aufgaben der Syntax. In *Syntax, vol. I*, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, & Theo Vennemann, 1-88. Berlin: de Gruyter.

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.

Sternefeld, Wolfgang. 1985. On Case and Binding Theory. In *Studies in German Grammar*, ed. Jindřich Toman, 231-285. Dordrecht: Foris.

Sternefeld, Wolfgang. 1985a. Deutsch ohne grammatische Funktionen. *Linguistische Berichte* 99:394-439.

Sternefeld, Wolfgang. 1990. Scrambling and Minimality. In *Scrambling and Barriers*, ed. Günther Grewendorf & Wolfgang Sternefeld, 239-257. Amsterdam: Benjamins.

Sternefeld, Wolfgang. 1991. *Syntaktische Grenzen*. Opladen: Westdeutscher Verlag.

Sternefeld, Wolfgang. 1991a. Chain Formation, Reanalysis, and the Economy of Levels. In *Derivation and Representation in the Theory of Grammar*, ed. Hubert Haider & Klaus Netter, 75-143. Dordrecht: Kluwer.

Sternefeld, Wolfgang. 1992. Transformationstypologie und Strukturelle Hierarchie. Ms., Universität Tübingen.

Sternefeld, Wolfgang. 1994. *Lecture Notes on Syntax*. Copenhagen: 6th European Summerschool of LLI.

Sternefeld, Wolfgang. 1995. Voice Phrases and Their Specifiers. Ms., Universität Tübingen. (*SfS-Report* 05/95).

Sternefeld, Wolfgang. 1996. A Minimalist Semantics for Questions. Ms., Universität Tübingen.

Sternefeld, Wolfgang. 1997. Comparing Reference Sets. In *Economy in Linguistic Theory*, ed. Manfred Bierwisch, Hans-Martin Gärtner, & Chris Wilder, 81-114. Berlin: Akademieverlag.

Sternefeld, Wolfgang. 1997a. The Semantics of Reconstruction and Connectivity. Ms., Universität Tübingen. (Arbeitspapiere des SFB 340 no. 97.)

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

Stowell, Tim. 1991. Small Clause Restructuring. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin, 182-218. Cambridge, Mass.: MIT Press.

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

Takano, Yuji. 1993. Minimalism and Proper Binding. Ms., University of California, Irvine.

Takano, Yuji. 1995. Predicate Fronting and Internal Subjects. *Linguistic Inquiry* 26:327-340.

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

Tappe, Thilo. 1989. A Note on Split Topicalization in German. In *Syntactic Phrase Structure Phenomena in Noun Phrases and Sentences*, ed. Christa Bhatt, Elisabeth Löbel, & Claudia Schmidt, 159-179. Amsterdam: Benjamins.

Thiersch, Craig. 1978. *Topics in German Syntax*. Doctoral dissertation, MIT, Cambridge, Mass.

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

Toman, Jindřich. 1986. Cliticization from NP in Czech. In *The Syntax of Pronominal Clitics* (Syntax and Semantics 19), ed. Hagit Borer, 123-145. Orlando etc.: Academic Press.

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

Torrego, Esther. 1985. On Empty Categories in Nominals. Ms., University of Boston.

Travis, Lisa. 1984. *Parameters and Effects of Word Order Variation*. Doctoral dissertation, MIT, Cambridge, Mass.

Travis, Lisa. 1991. Parameters of Phrase Structure and Verb-Second Phenomena. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin, 393-364. Cambridge, Mass.: MIT Press.

Truckenbrodt, Hubert. 1992. Fleeing the VP. Ms., MIT, Cambridge, Mass.

Truckenbrodt, Hubert. 1994. Towards a Prosodic Theory of Extraposition. Ms., MIT. Cambridge, Mass.

Vikner, Sten. 1990. *Verb Movement and the Licensing of NP Positions in the Germanic Languages*. Doctoral dissertation, Université de Genève.

Vikner, Sten. 1994. Scandinavian Object Shift and West Germanic Scrambling. In *Studies on Scrambling*, ed. Norbert Corver & Henk van Riemsdijk, 487-517. Berlin: Mouton de Gruyter.

Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford: Oxford University Press.

Wahba, Wafaa Abdel-Faheem Batran. 1992. LF Movement in Iraqi Arabic. In *Logical Structure and Linguistic Structure*, ed. C.-T. James Huang & Robert May, 253-276. Dordrecht: Kluwer.

Watanabe, Akira. 1992. Wh-in-Situ, Subjacency, and Chain Formation. Ms., MIT, Cambridge, Mass. (*MIT Occasional Papers in Linguistics* 2, Cambridge, Mass.: MITWPL.)

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

Webelhuth, Gert. 1987. Eine universale Scrambling-Theorie. Paper presented at the annual meeting of the DGfS, Augsburg.

Webelhuth, Gert. 1988. A Universal Theory of Scrambling. *Proceedings of the Conference on Scandinavian Linguistics* 10:284-298.
Webelhuth, Gert. 1992. *Principles and Parameters of Syntactic Saturation*. Oxford: Oxford University Press.
Wexler, Ken, & Peter Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge, Mass.: MIT Press.
Wilder, Chris. 1989. *The Syntax of German Infinitives*. Doctoral dissertation, University College London.
Wilder, Chris. 1996. Rightward Movement as Leftward Deletion. In *On Extraction and Extraposition in German*, ed. Uli Lutz & Jürgen Pafel, 273-309. Amsterdam: Benjamins.
Wilder, Chris, & Damir Čavar. 1994. Word Order Varation, Verb Movement and Economy Principles. *Studia Linguistica* 48:46-86.
Williams, Edwin. 1974. *Rule Ordering in Syntax*. Doctoral dissertation, MIT, Cambridge, Mass.
Williams, Edwin. 1983. Against Small Clauses. *Linguistic Inquiry* 14:287-308.
Zaenen, Annie. 1980. *Extraction Rules in Icelandic*. Doctoral dissertation, Harvard University, Cambridge, Mass.
Zwart, Jan-Wouter. 1993. *Dutch Syntax. A Minimalist Approach*. Doctoral dissertation, Rijksuniversiteit Groningen.
Zwart, Jan-Wouter. 1994. ‘Shortest Move’ vs. ‘Fewest Steps’. *Groninger Arbeiten zur Germanistischen Linguistik* 37:277-290.

Index